

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 1/20, 9/88, 9/90, 15/00, C12P
21/06, C07K 1/00, C07H 21/04

(21) International Application Number:

PCT/US98/20120

(11) International Publication Number: WO 99/15624

(43) International Publication Date: 1 April 1999 (01.04.99)

(74) Agent: McGURL, Barry, F.; Christensen, O'Connor, Johnson

US

••

(22) International Filing Date: 25 September 1998 (25.09.98)

25 September 1997 (25.09.97)

(63) Related by Continuation (CON) or Continuation-in-Part

(CIP) to Earlier Application
US 08/937,540 (CIP)
Filed on 25 September 1997 (25.09.97)

(71) Applicant (for all designated States except US): WASHING-TON STATE UNIVERSITY RESEARCH FOUNDATION [US/US]; N.E. 1615 Eastgate Boulevard, Pullman, WA 99164-1802 (US).

(72) Inventors; and

(30) Priority Data:

08/937,540

(75) Inventors/Applicants (for US only): CROTEAU, Rodney, Bruce [US/US]; 1835 N.E. Valley Road, Pullman, WA 99163 (US). WISE, Mitchell, Lynn [US/US]; 425A N.W. North Street, Pullman, WA 99163 (US). KATAHIRA, Eva, Joy [US/US]; 1720 N.E. Wheatland Drive, Pullman, WA 99163 (US). SAVAGE, Thomas, Jefferson [US/NZ]; 16 Cranbrook Avenue, Christchurch 5 (NZ).

(74) Agent: McGURL, Barry, F.; Christensen, O'Connor, Johnson & Kindness PLLC, Suite 2800, 1420 Fifth Avenue, Seattle, WA 98101 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: MONOTERPENE SYNTHASES FROM COMMON SAGE (SALVIA OFFICINALIS)

(57) Abstract

cDNAs encoding (+)-bornyl diphosphate synthase, 1-8-cineole synthase and (+)-sabinene synthase from common sage (Salvia officinalis) have been isolated and sequenced, and the corresponding amino acid sequences have been determined. Replicable recombinant cloning vehicles, host cells and the recombinant methods of making or enhanced expression of the enzymes in plants in order to increase the production of monoterpenoids or their products are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

4.1	A 16	cc	0	1.0	Tth	OT.	Clausaia	
AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia	
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia	
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal	
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland	
ΑZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad	
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo	
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan	
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan	
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey	
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago	
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine	
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda	
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America	
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan	
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam	
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia	
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe	
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand			
CM	Cameroon		Republic of Korea	PL	Poland			
CN	China	KR	Republic of Korea	PT	Portugal			
CU	Cuba	ΚZ	Kazakstan	RO	Romania			
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation			
DE	Germany	LI	Liechtenstein	SD	Sudan			
DK	Denmark	LK	Sri Lanka	SE	Sweden			
EE	Estonia	LR	Liberia	SG	Singapore			

Monoterpene Synthases From Common Sage (Salvia officinalis)

This invention was supported in part by grant numbers GM-31354 and DE-F603-9620212. The government has certain rights in the invention.

Field of the Invention

5

10

15

20

25

The present invention relates to nucleic acid sequences which code for monoterpene synthases (cyclases) from common sage (Salvia officinalis), and to vectors containing the sequences, host cells containing the sequences and methods of producing recombinant monoterpene synthases and their mutants.

Background of the Invention

The cyclization of the universal precursor geranyl diphosphate (GPP) to form monocyclic and bicyclic monoterpenes is catalyzed by a group of enzymes termed monoterpene synthases (or cyclases). The biochemical transformation of GPP to cyclic products has been investigated using enzymes from a variety of plants, including both angiosperms (Croteau, R., Chem. Rev. 87:929-954, 1987) and gymnosperms (Lewinsohn et al., Arch. Biochem. Biophys. 293:167-173, 1992, Savage et al., J. Biol. Chem. 269:4012-4020, 1994; Savage et al., Arch. Biochem. Biophys. 320:257-265, 1995). A mechanistic paradigm for these transformations is well established (Croteau, R., Chem. Rev. 87:929-954, 1987; Wise, M. L., and Croteau, R., in Comprehensive Natural Products Chemistry: Isoprenoids (Cane, D.E., ed) Vol. 2 (in press), Elsevier Science, Oxford, 1998). In summary, geranyl diphosphate is initially ionized and isomerized to form either 3R- or 3S-linalyl diphosphate, depending on the particular enzyme, which is converted to the α -terpinyl cation as a central intermediate. Further transformations of this reactive intermediate may be effected by additional intramolecular electrophilic additions, hydride shifts or other rearrangements before termination of the sequence by deprotonation of the final

cation or capture by an external nucleophile, such as a hydroxyl ion or the diphosphate group. Although the fate of the substrate has been well characterized in numerous monoterpene cyclization reactions, the molecular mechanisms by which the enzymes effect these transformations is still poorly understood.

5

10

15

20

25

30

35

Culinary sage (Salvia officinalis) produces a number of monoterpenes, including (+)- and (-)- α -pinene, (+)- and (-)- β -pinene, (+)- and (-)-camphene, (+)-sabinene, (+)- and (-)-limonene, myrcene, 1,8-cineole, and (+)-bornyl diphosphate (Croteau, R., Chem. Rev. 87:929-954, 1987). Because sage produces this broad range of acyclic, monocyclic and bicyclic monoterpenes, including several olefin isomers, a cyclic ether and a diphosphate ester, this plant has provided an ideal system for the study of a variety of biosynthetic enzymes, all of which utilize the same substrate but produce different products by variations on a single reaction mechanism (Croteau, R., Chem. Rev. 87:929-954, 1987; Wise, M. L., and Croteau, R., in Comprehensive Natural Products Chemistry: Isoprenoids (Cane, D.E., ed) Vol. 2 (in These monoterpene synthases include press), Elsevier Science, Oxford, 1998). (+)-bornyl diphosphate synthase (the enzyme producing the precursor of (+)-camphor) (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:523-532, 1979), 1,8-cineole synthase (Croteau et al, Arch. Biochem. Biophys. 309:184-192, 1994), (+)-sabinene synthase (the enzyme producing the precursor of (-)-3-isothujone) (Croteau, R., in Recent Developments in Flavor and Fragrance Chemistry (Hopp, R., and Mori, K., eds), pp. 263-273, VCH, Weinheim, Germany, 1992; Croteau, R., in Flavor Precursors: Thermal and Enzymatic Conversions (Teranishi, R., Takeoka, G. R., and Guntert, M., eds), American Chemical Society Symposium Series, No. 490, pp. 8-20, Washington, DC, 1992), and several pinene synthases (Gambliel, H., and Croteau, R., J. Biol. Chem. 257:2335-2342, 1982; Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984; Wagschal et al., Arch. Biochem. Biophys. 308:477-487, 1994; Pyun et al., Arch. Biochem. Biophys. 308:488-496, 1994).

As is typical of monoterpene cyclases, many of these enzymes from sage generate multiple products from geranyl diphosphate (Wise, M. L., and Croteau, R., in Comprehensive Natural Products Chemistry: Isoprenoids (Cane, D. E., ed) Vol. 2 (in press), Elsevier Science, Oxford, 1998; Wagschal et al., Tetrahedron 47:5933-5944, 1991). For example, investigations with the partially purified native enzymes have suggested that a single enzyme, termed (+)-pinene synthase (cyclase I), is responsible for the synthesis of both (+)- α -pinene and (+)-camphene, with lesser

5

10

15

20

25

30

35

amounts of (+)-limonene and myrcene, whereas a second enzyme, (-)-pinene synthase (cyclase II), has been shown to produce (-)-α-pinene, (-)-β-pinene and (-)-camphene, with minor amounts of (-)-limonene, terpinolene and myrcene (Gambliel, H., and Croteau, R., J. Biol. Chem. 257:2335-2342, 1982; Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984).

More recently, a third synthase from sage, termed cyclase III, has been described which produces a mixture of (+)- α -pinene and (+)- β -pinene, along with minor amounts of myrcene (Wagschal et al., Arch. Biochem. Biophys. 308:477-487, 1994; Pyun et al., Arch. Biochem. Biophys. 308:488-496, 1994). Evidence that these reactions are catalyzed by individual, multifunctional enzymes is provided by copurification and differential inhibition studies (Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984), as well as by isotopically sensitive branching experiments (Wagschal et al., Arch. Biochem. Biophys. 308:477-487, 1994; Wagschal et al., Tetrahedron 47:5933-5944, 1991; Croteau et al., Biochemistry 26:5383-5389, 1987). In spite of considerable effort, the (+)-pinene synthase has never been chromatographically separated from the aforementioned (+)-bornyl diphosphate synthase, suggesting that (+)-bornyl diphosphate synthase and (+)-pinene synthase might, in fact, be a single, multifunctional enzyme (McGeady, P., and Croteau, R., Arch. Biochem. Biophys. 317:149-155, 1995). Similarly, the (-)-pinene synthase has never been fully resolved from 1,8-cineole synthase, although, in this case, stereochemical considerations indicate that the two are distinct enzyme species (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994; Croteau et al., J. Biol. Chem. 264:2075-2080, 1989).

The unusual ability of the monoterpene synthases to synthesize multiple products from a single substrate requires the nomenclature of these enzymes to be based on the identity of the principal product synthesized by each enzyme. Thus, starting from the common precursor geranyl diphosphate, (+)-bornyl diphosphate synthase characteristically produces a mixture of monoterpenes of which at least 60% is (+)-bornyl diphosphate; 1,8-cineole synthase characteristically produces a mixture of monoterpenes of which at least 60% is 1,8-cineole and (+)-sabinene synthase characteristically produces a mixture of monoterpenes of which at least 60% is (+)-sabinene.

Summary of the Invention

In accordance with the foregoing, cDNAs encoding (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase from common sage (Salvia

5

10

15

20

25

30

35

officinalis) have been isolated and sequenced, and the corresponding amino acid sequences have been deduced. Accordingly, the present invention relates to isolated angiosperm DNA sequences which code for the expression of (+)-bornyl diphosphate synthase, such as the sequence designated SEQ ID No:1 which encodes (+)-bornyl diphosphate synthase from common sage (Salvia officinalis), for the expression of 1,8-cineole synthase, such as the sequence designated SEQ ID No:3, which encodes 1,8-cineole synthase from common sage (Salvia officinalis), and for the expression of (+)-sabinene synthase, such as the sequence designated SEQ ID No:5, which encodes the (+)-sabinene synthase from common sage (Salvia officinalis). In other aspects, the present invention is directed to replicable recombinant cloning vehicles comprising a nucleic acid sequence, e.g., a DNA sequence which codes for a (+)-bornyl diphosphate synthase, 1,8-cineole synthase or (+)-sabinene synthase, or for a base sequence sufficiently complementary to at least a portion of DNA or RNA encoding (+)-bornyl diphosphate synthase, 1,8-cineole synthase or (+)-sabinene synthase to enable hybridization therewith (e.g., antisense RNA or fragments of DNA complementary to a portion of DNA or RNA molecules encoding (+)-bornyl diphosphate synthase, 1,8-cineole synthase or (+)-sabinene synthase which are useful as polymerase chain reaction primers or as probes for any of the foregoing synthases or related genes). In yet other aspects of the invention, modified host cells are provided that have been transformed, transfected, infected and/or injected with a recombinant cloning vehicle and/or DNA sequence of the invention. Thus, the present invention provides for the recombinant expression of (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase, and the inventive concepts may be used to facilitate the production, isolation and purification of significant quantities of recombinant (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase (or of their primary enzyme products) for subsequent use, to obtain expression or enhanced expression of (+)-bornyl diphosphate synthase, 1,8cineole synthase and (+)-sabinene synthase in plants, microorganisms or animals, or may be otherwise employed in an environment where the regulation or expression of (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase is desired for the production of these synthases, or their enzyme products, or derivatives thereof.

Brief Description of the Drawings

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same becomes better understood by

-5-

reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:

FIGURE 1 shows the terpenoid biosynthetic reactions catalyzed by (+)-bornyl diphosphate synthase (1), (+)-sabinene synthase (2) and 1,8-cineole synthase (3).

Detailed Description of the Preferred Embodiment

As used herein, the terms "amino acid" and "amino acids" refer to all naturally occurring L- α -amino acids or their residues. The amino acids are identified by either the single-letter or three-letter designations:

	Asp	D	aspartic acid	Ile	I	isoleucine
10	Thr	T	threonine	Leu	L	leucine
	Ser	S	serine	Tyr	Y	tyrosine
	Glu	Е	glutamic acid	Phe	F	phenylalanine
	Pro	P	proline	His	H	histidine
	Gly	G	glycine	Lys	K	lysine
15	Ala	Α	alanine	Arg	R	arginine
	Cys	С	cysteine	Trp	W	tryptophan
	Val	V	valine	Gln	Q	glutamine
	Met	M	methionine	Asn	N	asparagine

5

20

25

30

35

As used herein, the term "nucleotide" means a monomeric unit of DNA or RNA containing a sugar moiety (pentose), a phosphate and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of pentose) and that combination of base and sugar is called a nucleoside. The base characterizes the nucleotide with the four bases of DNA being adenine ("A"), guanine ("G"), cytosine ("C") and thymine ("T"). Inosine ("I") is a synthetic base that can be used to substitute for any of the four, naturally-occurring bases (A, C, G or T). The four RNA bases are A,G,C and uracil ("U"). The nucleotide sequences described herein comprise a linear array of nucleotides connected by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

The term "angiosperm" refers to a group of plants that produce seeds that are enclosed within an ovary. An example of an angiosperm plant species is sage (Salvia officinalis).

The term "essential oil plant," or "essential oil plants," refers to a group of plant species that produce high levels of monoterpenoid and/or sesquiterpenoid and/or diterpenoid oils, and/or high levels of monoterpenoid and/or sesquiterpenoid and/or diterpenoid resins. The foregoing oils and/or resins account for greater than about

5

10

15

20

25

30

0.005% of the fresh weight of an essential oil plant that produces them. The essential oils and/or resins are more fully described, for example, in E. Guenther, The Essential Oils, Vols. I-VI, R.E. Krieger Publishing Co., Huntington N.Y., 1975, incorporated herein by reference. The essential oil plants include, but are not limited to:

Lamiaceae, including, but not limited to, the following species: Ocimum (basil), Lavandula (Lavender), Origanum (oregano), Mentha (mint), Salvia (sage), Rosmecinus (rosemary), Thymus (thyme), Satureja and Monarda.

Umbelliferae, including, but not limited to, the following species: Carum (caraway), Anethum (dill), feniculum (fennel) and Daucus (carrot).

Asteraceae (Compositae), including, but not limited to, the following species: Artemisia (tarragon, sage brush), Tanacetum (tansy).

Rutaceae (e.g., citrus plants); Rosaceae (e.g., roses); Myrtaceae (e.g., eucalyptus, Melaleuca); the Gramineae (e.g., Cymbopogon (citronella)); Geranaceae (Geranium) and certain conifers including Abies (e.g., Canadian balsam), Cedrus (cedar) and Thuja and Juniperus.

The range of essential oil plants is more fully set forth in E. Guenther, The Essential Oils, Vols. I-VI, R.E. Krieger Publishing Co., Huntington N.Y., 1975, which is incorporated herein by reference.

The term "percent identity" means the percentage of amino acids or nucleotides that occupy the same relative position when two amino acid sequences, or two nucleic acid sequences are aligned side by side.

The term "percent similarity" is a statistical measure of the degree of relatedness of two compared protein sequences. The percent similarity is calculated by a computer program that assigns a numerical value to each compared pair of amino acids based on chemical similarity (e.g., whether the compared amino acids are acidic, basic, hydrophobic, aromatic, etc.) and/or evolutionary distance as measured by the minimum number of base pair changes that would be required to convert a codon encoding one member of a pair of compared amino acids to a codon encoding the other member of the pair. Calculations are made after a best fit alignment of the two sequences have been made empirically by iterative comparison of all possible alignments. (Henikoff, S. and Henikoff, J.G., *Proc. Nat'l Acad Sci USA* 89:10915-10919, 1992).

"Oligonucleotide" refers to short length single or double stranded sequences of deoxyribonucleotides linked via phosphodiester bonds. The oligonucleotides are

chemically synthesized by known methods and purified, for example, on polyacrylamide gels.

The term "(+)-bornyl diphosphate synthase" is used herein to mean an enzyme capable of generating multiple monoterpenes from geranyl diphosphate. The principal and characteristic monoterpene synthesized by (+)-bornyl diphosphate synthase is bornyl pyrophosphate, which comprises at least 60% of the monoterpene mixture synthesized by (+)-bornyl diphosphate synthase from geranyl diphosphate.

5

10

15

20

25

30

35

The term "(+)-sabinene synthase" is used herein to mean an enzyme capable of generating multiple monoterpenes from geranyl diphosphate. The principal and characteristic monoterpene synthesized by (+)-sabinene synthase is sabinene, which comprises at least 60% of the monoterpene mixture synthesized by (+)-sabinene synthase from geranyl diphosphate.

The term "1,8-cineole synthase" is used herein to mean an enzyme capable of generating multiple monoterpenes from geranyl diphosphate. The principal and characteristic monoterpene synthesized by 1,8-cineole synthase is 1,8 cineole, which comprises at least 60% of the monoterpene mixture synthesized by 1,8-cineole synthase from geranyl diphosphate.

The terms "alteration", "amino acid sequence alteration", "variant" and "amino acid sequence variant" refer to (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase molecules with some differences in their amino acid sequences as compared to the corresponding, native, *i.e.*, naturally-occurring, synthases. Ordinarily, the variants will possess at least about 70% homology with the corresponding native synthases, and preferably, they will be at least about 80% homologous with the corresponding, native synthases. The amino acid sequence variants of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase falling within this invention possess substitutions, deletions, and/or insertions at certain positions. Sequence variants of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase may be used to attain desired enhanced or reduced enzymatic activity, modified regiochemistry or stereochemistry, or altered substrate utilization or product distribution.

Substitutional (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase variants are those that have at least one amino acid residue in the native synthase sequence removed and a different amino acid inserted in its place at the same position. The substitutions may be single, where only one amino acid in the molecule has been substituted, or they may be multiple, where two or more amino

acids have been substituted in the same molecule. Substantial changes in the activity of the (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase molecules may be obtained by substituting an amino acid with a side chain that is significantly different in charge and/or structure from that of the native amino acid. This type of substitution would be expected to affect the structure of the polypeptide backbone and/or the charge or hydrophobicity of the molecule in the area of the substitution.

Moderate changes in the activity of the (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase molecules would be expected by substituting an amino acid with a side chain that is similar in charge and/or structure to that of the native molecule. This type of substitution, referred to as a conservative substitution, would not be expected to substantially alter either the structure of the polypeptide backbone or the charge or hydrophobicity of the molecule in the area of the substitution.

10

15

20

25

30

35

Insertional (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase variants are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in the native synthase molecule. Immediately adjacent to an amino acid means connected to either the α -carboxy or α -amino functional group of the amino acid. The insertion may be one or more amino acids. Ordinarily, the insertion will consist of one or two conservative amino acids. Amino acids similar in charge and/or structure to the amino acids adjacent to the site of insertion are defined as conservative. Alternatively, this invention includes insertion of an amino acid with a charge and/or structure that is substantially different from the amino acids adjacent to the site of insertion.

Deletional variants are those where one or more amino acids in the native (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase molecules have been removed. Ordinarily, deletional variants will have one or two amino acids deleted in a particular region of the (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase molecule.

The terms "biological activity", "biologically active", "activity" and "active" refer to the ability of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase molecules to convert geranyl diphosphate to a group of monoterpenes, of which bornyl pyrophosphate is the principal and characteristic monoterpene synthesized by (+)-bornyl diphosphate synthase, sabinene is the principal and characteristic monoterpene synthesized by (+)-sabinene synthase and 1,8-cineole

is the principal and characteristic monoterpene synthesized by 1,8-cineole synthase. The monoterpenes produced by (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase are as measured in an enzyme activity assay, such as the assay described in Example 3. Amino acid sequence variants of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase may have desirable altered biological activity including, for example, altered reaction kinetics, substrate utilization product distribution or other characteristics such as regiochemistry and stereochemistry.

The terms "DNA sequence encoding", "DNA encoding" and "nucleic acid encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these deoxyribonucleotides determines the order of amino acids along the translated polypeptide chain. The DNA sequence thus codes for the amino acid sequence.

10

15

20

25

30

35

The terms "replicable expression vector" and "expression vector" refer to a piece of DNA, usually double-stranded, which may have inserted into it a piece of foreign DNA. Foreign DNA is defined as heterologous DNA, which is DNA not naturally found in the host. The vector is used to transport the foreign or heterologous DNA into a suitable host cell. Once in the host cell, the vector can replicate independently of or coincidental with the host chromosomal DNA, and several copies of the vector and its inserted (foreign) DNA may be generated. In addition, the vector contains the necessary elements that permit translating the foreign DNA into a polypeptide. Many molecules of the polypeptide encoded by the foreign DNA can thus be rapidly synthesized.

The terms "transformed host cell," "transformed" and "transformation" refer to the introduction of DNA into a cell. The cell is termed a "host cell", and it may be a prokaryotic or a eukaryotic cell. Typical prokaryotic host cells include various strains of *E. coli*. Typical eukaryotic host cells are plant cells, such as maize cells, yeast cells, insect cells or animal cells. The introduced DNA is usually in the form of a vector containing an inserted piece of DNA. The introduced DNA sequence may be from the same species as the host cell or from a different species from the host cell, or it may be a hybrid DNA sequence, containing some foreign DNA and some DNA derived from the host species.

The abbreviation "SSC" refers to a buffer used in nucleic acid hybridization solutions. One liter of the 20X (twenty times concentrate) stock SSC buffer solution (pH 7.0) contains 175.3 g sodium chloride and 88.2 g sodium citrate.

The following abbreviations are used herein: bp(s), base pair(s); DEAE, O-diethylaminoethyl; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GC, gas chromatography; IPTG, isopropyl-β-D-thiogalactopyranoside; LB, Luria-Bertani; Mopso, 3-(N-morpholino)-2-hydroxypropane-sulfonic acid; MS, mass spectrum/spectrometry; nt(s), nucleotide(s); ORF, open reading frame; PCR, polymerase chain reaction; PVDF, polyvinylidenedifluoride; SDS, sodium dodecyl sulfate; SBS, sage bornyl diphosphate synthase; SCS, sage 1,8-cineole synthase; SSS, sage sabinene synthase; Tris, tris(hydroxymethyl) aminomethane; UV, ultraviolet.

5

10

15

20

25

30

35

In accordance with the present invention, cDNAs encoding (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase were isolated and sequenced in the following manner. An homology-based PCR strategy was utilized that is a modified version of a published, homology-based PCR strategy (Steele et al., Proc. Natl. Acad. Sci. USA 92:4164-4168, 1995). A comparison was made of the deduced amino acid sequences of cDNAs encoding mechanisticallyrelated, but phylogenetically diverse, enzymes involved in terpenoid biosynthesis (Colby et al., J. Biol. Chem. 268:23016-23024, 1993; Yuba et al., Arch. Biochem. Biophys. 332:280-287, 1996; Bohlmann et al., J. Biol. Chem. 272:21784-21792, 1997). Three conserved regions of sequence were identified that appeared to be useful for the design of degenerate PCR primers. Two of these primers ultimately amplified a 600 bp fragment using cDNA from a phagemid sage leaf library as target. Cloning and sequencing showed that the amplified 600 bp product comprised two distinct sequence groups, both of which showed similarity to sequences of cloned terpene synthases, but only one of which hybridized strongly to a 2 kb target upon northern blot analysis of sage leaf mRNA. This more efficient probe was utilized to screen the sage leaf cDNA library, from which 77 positive phagemids were purified. Size selection of the purified and in vivo excised clones yielded a subset of 44 with inserts >1.6 kb, and these were expressed in E. coli XL1-Blue and assayed for functional monoterpene synthase activity by monitoring the conversion of [1-3H]geranyl diphosphate to monoterpene olefins, oxygenated monoterpenes and monoterpenyl diphosphate esters.

Two cDNA clones, of which the clone designated 3C6 was more active in expression, yielded an enzyme in the corresponding bacterial extracts that produced principally bornyl diphosphate from geranyl diphosphate. This recombinant enzyme, designated SBS (sage bornyl diphosphate synthase), was presumed to represent the native (+)-bornyl diphosphate synthase of sage, one of the prominent enzymes of oil

gland extracts (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979) that produces the first dedicated intermediate in (+)-camphor biosynthesis (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:523-532, 1979; Croteau, R., and Karp, F., Arch. Biochem. Biophys. 184:77-86, 1977; Croteau et al., Arch. Biochem. Biophys. 188:182-193, 1978).

5

10

15

20

25

30

Four cDNA clones, of which clone 3B5 was apparently most active, expressed a synthase in bacterial extracts that converted geranyl diphosphate to 1,8-cineole as the major product. This acquisition was designated SCS (sage 1,8-cineole synthase) and considered to represent the native 1,8-cineole synthase, an enzyme for which the mechanism of cyclization has been studied in detail (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994; Croteau, R., and Karp, F., Arch. Biochem. Biophys. 179:257-265, 1977).

Two additional clones, of which clone 3F25 was the more active in expression, yielded E. coli extracts capable of transforming geranyl diphosphate to sabinene as the dominant olefin product. This acquisition was named SSS (sage sabinene synthase), with correspondence assigned to the native (+)-sabinene synthase that catalyzes the cyclization to the bicyclic olefin precursor of (-)-isothujone (Croteau, R. in Recent Developments in Flavor and Fragrance Chemistry (Hopp, R., and Mori, K., eds), pp. 263-273, VCH, Weinheim, Germany, 1992; Croteau, R., in Flavor Precursors: Thermal and Enzymatic Conversions (Teranishi, R., Takeoka, G.R., and Guntert, M., eds), American Chemical Society Symposium Series, No. 490, pp. 8-20, Washington, DC, 1992; Karp, F., and Croteau, R., Arch. Biochem. Biophys. 216:616-624, 1982).

From the DNA sequence of Clones 3C6 (SEQ ID No:1), 3B5 (SEQ ID No:3) and 3F25 (SEQ ID No:5) the corresponding amino acid sequences of (+)-bornyl diphosphate synthase (SBS) (SEQ ID No:2), 1,8-cineole synthase (SCS) (SEQ ID No:4) and (+)-sabinene synthase (SSS) (SEQ ID No:6), respectively, were deduced.

Additionally, sequencing of cDNA clones that hybridized to the 600bp prenyltransferase probe, but which did not express detectable monoterpene cyclase activity in bacterial extracts, revealed an additional clone, designated 3F5 (SEQ ID No:7). The DNA sequence of clone 3F5 is similar to the sequences of clones 3C6, 3B5 and 3F25, and appears to represent a novel monoterpene cyclase clone. Clone 3F5 contains a premature, translational stop codon, consequently clone 3F5 does not encode a functional monoterpene cyclase.

The isolation of cDNAs encoding (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase permits the development of efficient expression systems for these functional enzymes; provides useful tools for examining the developmental regulation of monoterpene biosynthesis; permits investigation of the reaction mechanism(s) of these unusual, multiproduct enzymes, and permits the isolation of other (+)-bornyl diphosphate synthases, (+)-sabinene synthases. The isolation of the (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase cDNAs also permits the transformation of a wide range of organisms in order to introduce monoterpene biosynthesis de novo, or to modify endogenous monoterpene biosynthesis.

5

10

15

20

25

30

35

Although the (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase proteins set forth in SEQ ID Nos:2, 4 and 6, respectively, direct the enzymes to plastids, substitution of the targeting sequence of each of these enzymes (SEQ ID No:2, amino acids 1 to 56; SEQ ID No:4, amino acids 1 to 58; SEQ ID No:6, amino acids 1 to 53) with other transport sequences well known in the art (see, e.g., von Heijne et al., Eur. J. Biochem. 180:535-545, 1989; Stryer, Biochemistry, W.H. Freeman and Company, New York, NY, p. 769 [1988]) may be employed to direct the (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase to other cellular or extracellular locations.

In addition to the native (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase amino acid sequences of SEQ ID No:2, SEQ ID No:4 and SEQ ID No:6, respectively, sequence variants produced by deletions, substitutions, mutations and/or insertions are intended to be within the scope of the invention except insofar as limited by the prior art. The (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase amino acid sequence variants of this invention may be constructed by mutating the DNA sequences that encode the wild-type synthases, such as by using techniques commonly referred to as site-directed mutagenesis. Various polymerase chain reaction (PCR) methods, now well known in the field, such as a two primer system like the Transformer Site-Directed Mutagenesis kit from Clontech, may be employed for this purpose.

Following denaturation of the target plasmid in this system, two primers are simultaneously annealed to the plasmid; one of these primers contains the desired site-directed mutation, the other contains a mutation at another point in the plasmid resulting in elimination of a restriction site. Second strand synthesis is then carried out, tightly linking these two mutations, and the resulting plasmids are transformed

into a mutS strain of E. coli. Plasmid DNA is isolated from the transformed bacteria, restricted with the relevant restriction enzyme (thereby linearizing the unmutated plasmids), and then retransformed into E. coli. This system allows for generation of mutations directly in an expression plasmid, without the necessity of subcloning or generation of single-stranded phagemids. The tight linkage of the two mutations and the subsequent linearization of unmutated plasmids results in high mutation efficiency and allows minimal screening. Following synthesis of the initial restriction site primer, this method requires the use of only one new primer type per mutation site. Rather than prepare each positional mutant separately, a set of "designed degenerate" oligonucleotide primers can be synthesized in order to introduce all of the desired mutations at a given site simultaneously. Transformants can be screened by sequencing the plasmid DNA through the mutagenized region to identify and sort Each mutant DNA can then be restricted and analyzed by mutant clones. electrophoresis on Mutation Detection Enhancement gel (J.T. Baker) to confirm that no other alterations in the sequence have occurred (by band shift comparison to the unmutagenized control).

5

10

15

20

25

30

The verified mutant duplexes in the pET (or other) overexpression vector can be employed to transform E. coli such as strain E. coli BL21(DE3)pLysS, for high level production of the mutant protein, and purification by standard protocols. The method of FAB-MS mapping can be employed to rapidly check the fidelity of mutant expression. This technique provides for sequencing segments throughout the whole protein and provides the necessary confidence in the sequence assignment. In a mapping experiment of this type, protein is digested with a protease (the choice will depend on the specific region to be modified since this segment is of prime interest and the remaining map should be identical to the map of unmutagenized protein). The set of cleavage fragments is fractionated by microbore HPLC (reversed phase or ion exchange, again depending on the specific region to be modified) to provide several peptides in each fraction, and the molecular weights of the peptides are determined by FAB-MS. The masses are then compared to the molecular weights of peptides expected from the digestion of the predicted sequence, and the correctness of the Since this mutagenesis approach to protein sequence quickly ascertained. modification is directed, sequencing of the altered peptide should not be necessary if the MS agrees with prediction. If necessary to verify a changed residue, CAD-tandem MS/MS can be employed to sequence the peptides of the mixture in question, or the

target peptide purified for subtractive Edman degradation or carboxypeptidase Y digestion depending on the location of the modification.

5

10

15

20

25

30

35

In the design of a particular site directed mutagenesis, it is generally desirable to first make a non-conservative substitution (e.g., Ala for Cys, His or Glu) and determine if activity is greatly impaired as a consequence. The properties of the mutagenized protein are then examined with particular attention to the kinetic parameters of K_m and k_{cat} as sensitive indicators of altered function, from which changes in binding and/or catalysis per se may be deduced by comparison to the native enzyme. If the residue is by this means demonstrated to be important by activity impairment, or knockout, then conservative substitutions can be made, such as Asp for Glu to alter side chain length, Ser for Cys, or Arg for His. For hydrophobic segments, it is largely size that is usefully altered, although aromatics can also be substituted for alkyl side chains. Changes in the normal product distribution can indicate which step(s) of the reaction sequence have been altered by the mutation. Modification of the hydrophobic pocket can be employed to change binding conformations for substrates and result in altered regiochemistry and/or stereochemistry.

Other site directed mutagenesis techniques may also be employed with the nucleotide sequences of the invention. For example, restriction endonuclease digestion of DNA followed by ligation may be used to generate deletion variants of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase, as described in section 15.3 of Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, New York, NY [1989]). A similar strategy may be used to construct insertion variants, as described in section 15.3 of Sambrook et al., supra.

Oligonucleotide-directed mutagenesis may also be employed for preparing substitution variants of this invention. It may also be used to conveniently prepare the deletion and insertion variants of this invention. This technique is well known in the art as described by Adelman et al. (DNA 2:183 [1983]). Generally, oligonucleotides of at least 25 nucleotides in length are used to insert, delete or substitute two or more nucleotides in the (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase molecule. An optimal oligonucleotide will have 12 to 15 perfectly matched nucleotides on either side of the nucleotides coding for the mutation. To mutagenize wild-type (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase, the oligonucleotide is annealed to the single-stranded DNA template

molecule under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of *E. coli* DNA polymerase I, is then added. This enzyme uses the oligonucleotide as a primer to complete the synthesis of the mutation-bearing strand of DNA. Thus, a heteroduplex molecule is formed such that one strand of DNA encodes the wild-type synthase inserted in the vector, and the second strand of DNA encodes the mutated form of the synthase inserted into the same vector. This heteroduplex molecule is then transformed into a suitable host cell.

5

10

15

20

25

30

35

Mutants with more than one amino acid substituted may be generated in one of several ways. If the amino acids are located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If however, the amino acids are located some distance from each other (separated by more than ten amino acids, for example) it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed. In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions. An alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: wild-type (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase DNA is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more The oligonucleotide encoding the additional desired amino acid mutations. substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. resultant DNA can be used as a template in a third round of mutagenesis, and so on.

A gene encoding (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase may be incorporated into any organism (intact plant, animal, microbe, etc.), or cell culture derived therefrom, that produces geranyl diphosphate. A (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase gene may be introduced into any organism for a variety of purposes including, but not limited to: production of (+)-bornyl diphosphate synthase, (+)-sabinene synthase or

1,8-cineole synthase, or their products; production or modification of flavor and aroma properties; improvement of defense capability, and the alteration of other ecological interactions mediated by bornyl pyrophosphate, sabinene, 1,8-cineole, or their derivatives.

5

10

15

20

25

30

35

Eukaryotic expression systems may be utilized for the production of (+)bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase since they are capable of carrying out any required posttranslational modifications and of directing the enzymes to the proper membrane location. A representative eukaryotic expression system for this purpose uses the recombinant baculovirus, Autographa californica nuclear polyhedrosis virus (AcNPV; M.D. Summers and G.E. Smith, A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures [1986]; Luckow et al., Bio-technology 6:47-55 [1987]) for expression of the terpenoid synthases of the invention. Infection of insect cells (such as cells of the species Spodoptera frugiperda) with the recombinant baculoviruses allows for the production of large amounts of the monoterpenoid synthase proteins. In addition, the baculovirus system has other important advantages for the production of recombinant (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase. For example, baculoviruses do not infect humans and can therefore be safely handled in large quantities. In the baculovirus system, a DNA construct is prepared including a DNA segment encoding (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase and a vector. The vector may comprise the polyhedron gene promoter region of a baculovirus, the baculovirus flanking sequences necessary for proper cross-over during recombination (the flanking sequences comprise about 200-300 base pairs adjacent to the promoter sequence) and a bacterial origin of replication which permits the construct to replicate in bacteria. constructed so that (i) the DNA segment is placed adjacent (or operably linked or "downstream" or "under the control of") to the polyhedron gene promoter and (ii) the promoter/monoterpene synthase combination is flanked on both sides by 200-300 base pairs of baculovirus DNA (the flanking sequences).

To produce the monoterpene synthase DNA construct, a cDNA clone encoding the full length (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase is obtained using methods such as those described herein. The DNA construct is contacted in a host cell with baculovirus DNA of an appropriate baculovirus (that is, of the same species of baculovirus as the promoter encoded in the construct) under conditions such that recombination is effected. The resulting

recombinant baculoviruses encode the full (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase. For example, an insect host cell can be cotransfected or transfected separately with the DNA construct and a functional baculovirus. Resulting recombinant baculoviruses can then be isolated and used to infect cells to effect production of the monoterpene synthase. Host insect cells include, for example, *Spodoptera frugiperda* cells, that are capable of producing a baculovirus-expressed monoterpene synthase. Insect host cells infected with a recombinant baculovirus of the present invention are then cultured under conditions allowing expression of the baculovirus-encoded (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase thus produced are then extracted from the cells using methods known in the art.

5

10

15

20

25

30

35

Other eukaryotic microbes such as yeasts may also be used to practice this invention. The baker's yeast Saccharomyces cerevisiae, is a commonly used yeast, although several other strains are available. The plasmid YRp7 (Stinchcomb et al., Nature 282:39 [1979]; Kingsman et al., Gene 7:141 [1979]; Tschemper et al., Gene 10:157 [1980]) is commonly used as an expression vector in Saccharomyces. This plasmid contains the trp1 gene that provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, such as strains ATCC No. 44,076 and PEP4-1 (Jones, Genetics 85:12 [1977]). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan. Yeast host cells are generally transformed using the polyethylene glycol method, as described by Hinnen (Proc. Natl. Acad. Sci. USA 75:1929 [1978]). Additional yeast transformation protocols are set forth in Gietz et al., N.A.R. 20(17):1425, 1992; Reeves et al., FEMS 99:193-197, 1992.

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., J. Biol. Chem. 255:2073 [1980]) or other glycolytic enzymes (Hess et al., J. Adv. Enzyme Reg. 7:149 [1968]; Holland et al., Biochemistry 17:4900 [1978]), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In the construction of suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed

to provide polyadenylation of the mRNA and termination. Other promoters that have the additional advantage of transcription controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing yeast-compatible promoter, origin of replication and termination sequences is suitable.

10

15

20

25

30

35

Cell cultures derived from multicellular organisms, such as plants, may be used as hosts to practice this invention. Transgenic plants can be obtained, for example, by transferring plasmids that encode (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase and a selectable marker gene, e.g., the kan gene encoding resistance to kanamycin, into Agrobacterium tumifaciens containing a helper Ti plasmid as described in Hoeckema et al., Nature 303:179-181 [1983] and culturing the Agrobacterium cells with leaf slices of the plant to be transformed as described by An et al., Plant Physiology 81:301-305 [1986]. Transformation of cultured plant host cells is normally accomplished through Agrobacterium tumifaciens, as described above. Cultures of mammalian host cells and other host cells that do not have rigid cell membrane barriers are usually transformed using the calcium phosphate method as originally described by Graham and Van der Eb (Virology 52:546 [1978]) and modified as described in sections 16.32-16.37 of Sambrook et al., supra. However, other methods for introducing DNA into cells such as Polybrene (Kawai and Nishizawa, Mol. Cell. Biol. 4:1172 [1984]), protoplast fusion (Schaffner, Proc. Natl. Acad. Sci. USA 77:2163 [1980]), electroporation (Neumann et al., EMBO J. 1:841 [1982]), and direct microinjection into nuclei Additionally, animal (Capecchi, Cell 22:479 [1980]) may also be used. transformation strategies are reviewed in Monastersky G.M. and Robl, J.M., Strategies in Transgenic Animal Science, ASM Press, Washington, D.C., 1995. Transformed plant calli may be selected through the selectable marker by growing the cells on a medium containing, e.g., kanamycin, and appropriate amounts of phytohormone such as naphthalene acetic acid and benzyladenine for callus and shoot induction. The plant cells may then be regenerated and the resulting plants transferred to soil using techniques well known to those skilled in the art.

In addition, a gene regulating (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase production can be incorporated into the plant along with a necessary promoter which is inducible. In the practice of this embodiment of

the invention, a promoter that only responds to a specific external or internal stimulus is fused to the target cDNA. Thus, the gene will not be transcribed except in response to the specific stimulus. As long as the gene is not being transcribed, its gene product is not produced.

5

10

15

20

25

30

35

An illustrative example of a responsive promoter system that can be used in the practice of this invention is the glutathione-S-transferase (GST) system in maize. GSTs are a family of enzymes that can detoxify a number of hydrophobic electrophilic compounds that often are used as pre-emergent herbicides (Weigand et al., Plant Molecular Biology 7:235-243 [1986]). Studies have shown that the GSTs are directly involved in causing this enhanced herbicide tolerance. This action is primarily mediated through a specific 1.1 kb mRNA transcription product. In short, maize has a naturally occurring quiescent gene already present that can respond to external stimuli and that can be induced to produce a gene product. This gene has previously been identified and cloned. Thus, in one embodiment of this invention, the promoter is removed from the GST responsive gene and attached to a (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase gene that previously has had its native promoter removed. This engineered gene is the combination of a promoter that responds to an external chemical stimulus and a gene responsible for successful production of (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase.

In addition to the methods described above, several methods are known in the art for transferring cloned DNA into a wide variety of plant species, including gymnosperms, angiosperms, monocots and dicots (see, e.g., Glick and Thompson, eds., Methods in Plant Molecular Biology, CRC Press, Boca Raton, Florida [1993]). Representative examples include electroporation-facilitated DNA uptake by protoplasts (Rhodes et al., Science 240(4849):204-207 [1988]); treatment of protoplasts with polyethylene glycol (Lyznik et al., Plant Molecular Biology 13:151-161 [1989]); and bombardment of cells with DNA laden microprojectiles (Klein et al., Plant Physiol. 91:440-444 [1989] and Boynton et al., Science 240(4858):1534-1538 [1988]). Additionally, plant transformation strategies and techniques are reviewed in Birch, R.G., Ann Rev Plant Phys Plant Mol Biol 48:297 (1997); Forester et al., Exp. Agric. 33:15-33 (1997). Minor variations make these technologies applicable to a broad range of plant species.

Each of these techniques has advantages and disadvantages. In each of the techniques, DNA from a plasmid is genetically engineered such that it contains not

only the gene of interest, but also selectable and screenable marker genes. selectable marker gene is used to select only those cells that have integrated copies of the plasmid (the construction is such that the gene of interest and the selectable and screenable genes are transferred as a unit). The screenable gene provides another check for the successful culturing of only those cells carrying the genes of interest. A commonly used selectable marker gene is neomycin phosphotransferase II (NPT II). This gene conveys resistance to kanamycin, a compound that can be added directly to the growth media on which the cells grow. Plant cells are normally susceptible to kanamycin and, as a result, die. The presence of the NPT II gene overcomes the effects of the kanamycin and each cell with this gene remains viable. selectable marker gene which can be employed in the practice of this invention is the gene which confers resistance to the herbicide glufosinate (Basta). A screenable gene commonly used is the \beta-glucuronidase gene (GUS). The presence of this gene is characterized using a histochemical reaction in which a sample of putatively transformed cells is treated with a GUS assay solution. After an appropriate incubation, the cells containing the GUS gene turn blue.

10

15

20

25

30

35

The plasmid containing one or more of these genes is introduced into either plant protoplasts or callus cells by any of the previously mentioned techniques. If the marker gene is a selectable gene, only those cells that have incorporated the DNA package survive under selection with the appropriate phytotoxic agent. Once the appropriate cells are identified and propagated, plants are regenerated. Progeny from the transformed plants must be tested to insure that the DNA package has been successfully integrated into the plant genome.

Mammalian host cells may also be used in the practice of the invention. Examples of suitable mammalian cell lines include monkey kidney CVI line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line 293S (Graham et al., *J. Gen. Virol.* 36:59 [1977]); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells (Urlab and Chasin, *Proc. Natl. Acad. Sci USA* 77:4216 [1980]); mouse sertoli cells (TM4, Mather, *Biol. Reprod.* 23:243 [1980]); monkey kidney cells (CVI-76, ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor cells (MMT 060562, ATCC CCL 51); rat hepatoma cells (HTC, MI.54, Baumann et al., *J. Cell Biol.* 85:1

[1980]); and TRI cells (Mather et al., Annals N.Y. Acad. Sci. 383:44 [1982]). Expression vectors for these cells ordinarily include (if necessary) DNA sequences for an origin of replication, a promoter located in front of the gene to be expressed, a ribosome binding site, an RNA splice site, a polyadenylation site, and a transcription terminator site.

5

10

15

20

25

30

35

Promoters used in mammalian expression vectors are often of viral origin. These viral promoters are commonly derived from polyoma virus, Adenovirus 2, and most frequently Simian Virus 40 (SV40). The SV40 virus contains two promoters that are termed the early and late promoters. These promoters are particularly useful because they are both easily obtained from the virus as one DNA fragment that also contains the viral origin of replication (Fiers et al., *Nature* 273:113 [1978]). Smaller or larger SV40 DNA fragments may also be used, provided they contain the approximately 250-bp sequence extending from the HindIII site toward the BglI site located in the viral origin of replication.

Alternatively, promoters that are naturally associated with the foreign gene (homologous promoters) may be used provided that they are compatible with the host cell line selected for transformation.

An origin of replication may be obtained from an exogenous source, such as SV40 or other virus (e.g., Polyoma, Adeno, VSV, BPV) and inserted into the cloning vector. Alternatively, the origin of replication may be provided by the host cell chromosomal replication mechanism. If the vector containing the foreign gene is integrated into the host cell chromosome, the latter is often sufficient.

The use of a secondary DNA coding sequence can enhance production levels of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase in transformed cell lines. The secondary coding sequence typically comprises the enzyme dihydrofolate reductase (DHFR). The wild-type form of DHFR is normally inhibited by the chemical methotrexate (MTX). The level of DHFR expression in a cell will vary depending on the amount of MTX added to the cultured host cells. An additional feature of DHFR that makes it particularly useful as a secondary sequence is that it can be used as a selection marker to identify transformed cells. Two forms of DHFR are available for use as secondary sequences, wild-type DHFR and MTX-resistant DHFR. The type of DHFR used in a particular host cell depends on whether the host cell is DHFR deficient (such that it either produces very low levels of DHFR endogenously, or it does not produce functional DHFR at all). DHFR-deficient cell lines such as the CHO cell line described by Urlaub and Chasin, supra, are

-22-

transformed with wild-type DHFR coding sequences. After transformation, these DHFR-deficient cell lines express functional DHFR and are capable of growing in a culture medium lacking the nutrients hypoxanthine, glycine and thymidine. Nontransformed cells will not survive in this medium.

5

10

15

20

25

30

35

The MTX-resistant form of DHFR can be used as a means of selecting for transformed host cells in those host cells that endogenously produce normal amounts of functional DHFR that is MTX sensitive. The CHO-Kl cell line (ATCC No. CL 61) possesses these characteristics, and is thus a useful cell line for this purpose. The addition of MTX to the cell culture medium will permit only those cells transformed with the DNA encoding the MTX-resistant DHFR to grow. The nontransformed cells will be unable to survive in this medium.

Prokaryotes may also be used as host cells for the initial cloning steps of this invention. They are particularly useful for rapid production of large amounts of DNA, for production of single-stranded DNA templates used for site-directed mutagenesis, for screening many mutants simultaneously, and for DNA sequencing of the mutants generated. Suitable prokaryotic host cells include E. coli K12 strain 94 (ATCC No. 31,446), E. coli strain W3110 (ATCC No. 27,325) E. coli X1776 (ATCC No. 31,537), and E. coli B; however many other strains of E. coli, such as HB101, JM101, NM522, NM538, NM539, and many other species and genera of prokaryotes including bacilli such as Bacillus subtilis, other enterobacteriaceae such as Salmonella typhimurium or Serratia marcesans, and various Pseudomonas species may all be used as hosts. Prokaryotic host cells or other host cells with rigid cell walls are preferably transformed using the calcium chloride method as described in section 1.82 Alternatively, electroporation may be used for of Sambrook et al., supra. transformation of these cells. Prokaryote transformation techniques are set forth in Dower, W. J., in Genetic Engineering, Principles and Methods, 12:275-296, Plenum Publishing Corp., 1990; Hanahan et al., Meth. Enxymol., 204:63, 1991.

As a representative example, cDNA sequences encoding (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase may be transferred to the (His)₆•Tag pET vector commercially available (from Novagen) for overexpression in *E. coli* as heterologous host. This pET expression plasmid has several advantages in high level heterologous expression systems. The desired cDNA insert is ligated in frame to plasmid vector sequences encoding six histidines followed by a highly specific protease recognition site (thrombin) that are joined to the amino terminus codon of the target protein. The histidine "block" of the expressed fusion

protein promotes very tight binding to immobilized metal ions and permits rapid purification of the recombinant protein by immobilized metal ion affinity chromatography. The histidine leader sequence is then cleaved at the specific proteolysis site by treatment of the purified protein with thrombin, and the (+)-bornyl diphosphate synthase, (+)-sabinene synthase or 1,8-cineole synthase again purified by immobilized metal ion affinity chromatography, this time using a shallower imidazole gradient to elute the recombinant synthases while leaving the histidine block still adsorbed. This overexpression-purification system has high capacity, excellent resolving power and is fast, and the chance of a contaminating *E. coli* protein exhibiting similar binding behavior (before and after thrombin proteolysis) is extremely small.

10

15

20

25

30

35

As will be apparent to those skilled in the art, any plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell may also be used in the practice of the invention. The vector usually has a replication site, marker genes that provide phenotypic selection in transformed cells, one or more promoters, and a polylinker region containing several restriction sites for insertion of foreign DNA. Plasmids typically used for transformation of *E. coli* include pBR322, pUC18, pUC19, pUC118, pUC119, and Bluescript M13, all of which are described in sections 1.12-1.20 of Sambrook et al., *supra*. However, many other suitable vectors are available as well. These vectors contain genes coding for ampicillin and/or tetracycline resistance which enables cells transformed with these vectors to grow in the presence of these antibiotics.

The promoters most commonly used in prokaryotic vectors include the β-lactamase (penicillinase) and lactose promoter systems (Chang et al. Nature 375:615 [1978]; Itakura et al., Science 198:1056 [1977]; Goeddel et al., Nature 281:544 [1979]) and a tryptophan (trp) promoter system (Goeddel et al., Nucl. Acids Res. 8:4057 [1980]; EPO Appl. Publ. No. 36,776), and the alkaline phosphatase systems. While these are the most commonly used, other microbial promoters have been utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to ligate them functionally into plasmid vectors (see Siebenlist et al., Cell 20:269 [1980]).

Many eukaryotic proteins normally secreted from the cell contain an endogenous secretion signal sequence as part of the amino acid sequence. Thus, proteins normally found in the cytoplasm can be targeted for secretion by linking a signal sequence to the protein. This is readily accomplished by ligating DNA

encoding a signal sequence to the 5' end of the DNA encoding the protein and then expressing this fusion protein in an appropriate host cell. The DNA encoding the signal sequence may be obtained as a restriction fragment from any gene encoding a protein with a signal sequence. Thus, prokaryotic, yeast, and eukaryotic signal sequences may be used herein, depending on the type of host cell utilized to practice the invention. The DNA and amino acid sequence encoding the signal sequence portion of several eukaryotic genes including, for example, human growth hormone, proinsulin, and proalbumin are known (see Stryer, *Biochemistry* W.H. Freeman and Company, New York, NY, p. 769 [1988]), and can be used as signal sequences in appropriate eukaryotic host cells. Yeast signal sequences, as for example acid phosphatase (Arima et al., *Nuc. leids Res.* 11:1657 [1983]), α-factor, alkaline phosphatase and invertase may be used to direct secretion from yeast host cells. Prokaryotic signal sequences from genes encoding, for example, LamB or OmpF (Wong et al., *Gene* 68:193 [1988]), MalE, PhoA, or beta-lactamase, as well as other genes, may be used to target proteins from prokaryotic cells into the culture medium.

5

10

15

20

25

30

35

As described above, the (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase amino terminal membrane insertion sequences reside at SEQ ID No:2, amino acids 1 through 56; SEQ ID No:4, amino acids 1 through 58; SEQ ID No:6, amino acids 1 through 53) and direct the enzymes to plastids. Alternative trafficking sequences from plants, animals and microbes can be employed in the practice of the invention to direct the gene product to the cytoplasm, endoplasmic reticulum, mitochondria or other cellular components, or to target the protein for export to the medium. These considerations apply to the overexpression of (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase, and to direction of expression within cells or intact organisms to permit gene product function in any desired location.

The construction of suitable vectors containing DNA encoding replication sequences, regulatory sequences, phenotypic selection genes and the monoterpene synthase DNA of interest are prepared using standard recombinant DNA procedures. Isolated plasmids and DNA fragments are cleaved, tailored, and ligated together in a specific order to generate the desired vectors, as is well known in the art (see, for example, Maniatis, *supra*, and Sambrook et al., *supra*).

As discussed above, (+)-bornyl diphosphate synthase, (+)-sabinene synthase and 1,8-cineole synthase variants are preferably produced by means of mutation(s) that are generated using the method of site-specific mutagenesis. This method

-25-

requires the synthesis and use of specific oligonucleotides that encode both the sequence of the desired mutation and a sufficient number of adjacent nucleotides to allow the oligonucleotide to stably hybridize to the DNA template.

The foregoing may be more fully understood in connection with the following representative examples, in which "Plasmids" are designated by a lower case p followed by an alphanumeric designation. The starting plasmids used in this invention are either commercially available, publicly available on an unrestricted basis, or can be constructed from such available plasmids using published procedures. In addition, other equivalent plasmids are known in the art and will be apparent to the ordinary artisan.

"Digestion", "cutting" or "cleaving" of DNA refers to catalytic cleavage of the DNA with an enzyme that acts only at particular locations in the DNA. These enzymes are called restriction endonucleases, and the site along the DNA sequence where each enzyme cleaves is called a restriction site. The restriction enzymes used in this invention are commercially available and are used according to the instructions supplied by the manufacturers. (See also sections 1.60-1.61 and sections 3.38-3.39 of Sambrook et al., supra.)

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the resulting DNA fragment on a polyacrylamide or an agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally. For example, see Lawn et al. (Nucleic Acids Res. 9:6103-6114 [1982]), and Goeddel et al. (Nucleic Acids Res., supra).

The following examples merely illustrate the best mode now contemplated for practicing the invention, but should not be construed to limit the invention. All literature citations herein are expressly incorporated by reference.

EXAMPLES

30 Example 1

5

10

15

20

25

35

cDNA Library Construction and Cloning of Monoterpene Synthases

cDNA Library Preparatiom, Sage plants (S. officinalis L.) were grown from seed as previously described (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979). Approximately 15 g of emerging sage leaves (shoot tips) from three-week-old plants were ground to a fine powder in liquid nitrogen and extracted

5

10

15

20

25

30

35

into buffer composed of 200 mM Tris-HCl (pH 8.5), 300 mM LiCl and 10 mM EDTA, and containing 1% (w/v) polyvinylpyrrolidone (M_r ~40,000). The high concentration of chloride salts and high pH were empirically optimized to maximize the yield of intact RNA, and polyvinylpyrrolidone was found to be essential to complex co-extracted oils, resins and phenolic substances that otherwise prevent RNA isolation. Total RNA thus extracted was prepared by precipitation with isopropanol, followed by CsCl density gradient centrifugation, as previously described (Lewinsohn et al., *Plant Mol. Biol. Rep.* 12:20-25, 1994). Poly(A)⁺ mRNA was isolated by chromatography on oligo(dT)-cellulose (Qiagen) and 6.3 µg of the resulting mRNA was used to construct a Σ ZAPII cDNA library according to the manufacturer's instructions (Stratagene).

PCR-Based Probe Generation and Library Screening Protein purification from sage, as the basis for cDNA isolation, has been of limited success (McGeady, P., and Croteau, R., Arch. Biochem. Biophys. 317:149-155, 1995) because of the number of synthases present and their similarity in physical properties (Alonso, W. R., and Croteau, R., in Methods in Plant Biochemistry (Enzymes of Secondary Metabolism) (Lea, P. J., ed) Vol. 9, pp. 239-260, Academic Press, New York, 1993), and thus far has not permitted a reverse genetic approach to cloning of any of the monoterpene synthases from this species. Consequently, a generic strategy for the homology-based PCR cloning of terpenoid synthases of higher plant origin was utilized (Steele et al., Proc. Natl. Acad. Sci. USA 92:4164-4168, 1995), in this instance by comparing monoterpene synthase cDNA sequences that were isolated from both angiosperms and gymnosperms (Colby et al., J. Biol. Chem. 268:23016-23024, 1993; Yuba et al., Arch. Biochem. Biophys. 332:280-287, 1996; Bohlmann et al., J. Biol. Chem. 272:21784-21792, 1997).

Three PCR oligonucleotide primers were synthesized based on the results of the monoterpene synthase homology comparison:

 $1F \quad 5'AA(G/A)AA(T/C)GA(G/A)(G/A)A(G/A)GGIGAITA(C/T)AA(G/A)GA-3'\\ (SEQ ID No:8)$

2F 5'-(T/C)TICA(G/A)(C/T)TITA(T/C)GA(G/A)GC-3' (SEQ ID No:9)

3R 5'-CT(A/G)GT(C/T)(G/A)AIGGI(C/A)T(G/A)AT(G/A)TACGT(C/T)-3' (SEQ ID No:10)

Each of the sense primers (1F and 2F) was used for PCR in combination with antisense primer (3R). Using purified sage leaf cDNA library phage as template (5 μ l at 1.5 x 10⁹ plaque forming units/ml), PCR was performed in a total volume of 50 μ l

containing 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 5 mM MgCl₂, 200 µM of each dNTP, 5 µM of each primer and 2.5 units of Taq polymerase (BRL or Life Sciences). The thermal cycler performed the following denaturation, annealing and amplification steps: denaturation at 94°C, 1 minute; annealing at 60°C, 1 minute; extension at 72°C, 3.5 minutes; 35 cycles with final extension at 72°C, five minutes. Analysis of the PCR reaction products by agarose gel electrophoresis (Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) indicated that only the combination of primers 2F and 3R amplified a discrete product of approximately 600 bp, which was gel purified, ligated into pT7BIue (Novagen), and transformed Plasmid DNA was prepared from 32 individual into E. coli NovaBlue cells. transformants and the inserts partially sequenced (DyeDeoxy Terminator Cycle Sequencing; Applied Biosystems) to reveal that two distinct "terpenoid synthase-like" sequences had been amplified in roughly comparable amounts (SEQ ID No:11 and SEQ ID No:12).

5

10

15

20

25

30

35

The relative ability of these two potential probes to hybridize with expressed genes was evaluated by RNA-DNA hybridization. Two samples of sage leaf mRNA isolated as above (3 µg each) were electrophoresed on 1% (w/v) agarose under denaturing conditions and blotted onto separate PVDF membranes using standard techniques (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). Each membrane was evaluated with ³²P labeled probe, generated from one or the other of the 600 bp fragments using random hexamer priming (Tabor et al., Current Protocols in Molecular Biology (Ausubel et al., eds), sections 3.5.9-3.5.10, John Wiley & Sons, New York, 1991), by standard hybridization and washing protocols (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). Autoradiography of the membrane revealed that both probes hybridized to a 2 kb transcript, although one probe (SEQ ID No:11) generated a significantly stronger signal (~10-fold) than the other (SEQ ID No:12). The probe generating the stronger signal (SEQ ID No:11) was subsequently employed to screen the cDNA library in an attempt to isolate full-length cDNA sequences encoding the corresponding terpene synthase.

UV-crosslinked nitrocellulose lifts containing 3-5 x 10^4 primary plaques (plated on *E. coli* XL1-Blue-MRF'), after pre-hybridization (in 1.25 x SSPE, 0.5 x Denhart's reagent, 9% formamide, 0.002% SDS, and 10 μ g/ml denatured *E. coli*

DNA, for 2 h at 42°C), were hybridized in the same medium with approximately 8 μCi of the ³²P labeled probe for 48 h. Filters were washed, first at room temperature (in 2 x SSC with 0.1% SDS), then at 55°C (in 1 x SSC with 0.1% SDS), and subsequently exposed to X-ray film at -70°C (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). Plaques yielding positive signals were purified through two additional rounds of hybridization. A total of 77 purified ΣΖΑΡ clones so isolated were excised in vivo to generate BluescriptII SK(-) phagmids and transformed into E. coli XLOLR cells according to the Stratagene protocol. The size of each cDNA insert was determined by PCR using T3 and T7 primers, and transformed clones containing an insert >1.6 kb were either expressed to assay for monoterpene synthase activity or sequenced at the 5'-terminus using the T3 promoter Bluescript plasmids expressing synthase activity in cell-free extracts of transformed E. coli (see Examples 2 and 3) were fully sequenced on both DNA strands by primer walking or by the method of nested deletions using exonuclease III and mung bean nuclease (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).

5

10

15

20

25

30

35

To improve functional expression and facilitate subsequent enzyme purification, each of the apparent full-length pBluescript clones that expressed monoterpene synthase activity was subcloned in frame into pGEX vectors (Pharmacia) using a convenient BamHI (SBC and SSS) or EcoRI (SCS) restriction site at the 5'-end, and the XhoI restriction site at the 3'-terminus. Fidelity in subcloning was confirmed by complete sequencing, and these plasmid constructs were expressed in E. coli XL1-Blue-MFR' cells.

Example 2

Expression of Monoterpene Synthase cDNAs in E. coli

The bluescript plasmids expressed in *E. coli* strain XL1-Blue were grown in 5 ml LB medium (Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., pp. 2.69-2.76, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989), supplemented with 100 μ g ampicillin/ml, to an $A_{600} = 0.5$ at 37°C with constant shaking, then induced with 1 to 3 mM IPTG. The cells were allowed an additional 4 h growth at 37°C before harvesting by centrifugation (2000 x g, 10 min) and lysis by sonication (Braun-Sonic 2000 with microprobe at maximum power for 15 seconds), on ice, in 50 mM Mopso buffer containing 10% glycerol, 10 mM MgCl₂

and 5 mM DTT (pH either 6.5 or 7.1, as appropriate). The sonicates were cleared by centrifugation (18,000 x g, 10 min) and the resulting supernatant was used as the enzyme source. The pGEX constructs in E. coli XLl-Blue-MFR' cells were similarly grown at 37°C to $A_{600} = 1.0$ to 1.5, then induced with 1 mM IPTG and incubated overnight at 20°C with constant shaking. The cells were then harvested and lysed, and the soluble supernatant prepared as before. Purification of the resulting fusion proteins was attempted using the glutathione-Sepharose affinity column according to the manufacturer's instructions (Pharmacia).

5

10

15

20

25

30

35

Of the three expressed monoterpene synthases (SBS, SCS and SSS), only one (SBS) bound to the matrix but, even in this case, affinity-based purification proved to Therefore, partial purification of the heterologously expressed be unreliable. synthases was achieved by ion-exchange chromatography on DEAE-cellulose (Whatman DE-52) using a 0-400 mM NaCl gradient. The partially purified preparations were desalted by repeated ultrafiltration and dilution using an Amicon Centriprep 30 concentrator (30 kDa cutoff) and the appropriate assay buffer. The pGEX-expressed fusion proteins were also subjected to gel permeation chromatography (Pharmacia FPLC system) using a Pharmacia XY 16 x 70 column packed with Superdex S-200 and equilibrated with the appropriate 50 mM Mopso buffer system. The column was developed at a flow rate of 0.3 ml per min and was calibrated using the Sigma MW-GF-200 molecular weight marker kit. Kav values of the recombinant enzymes were compared to the calibration standards to establish molecular weights (Cooper, T. G. The Tools of Biochemistry, John Wiley & Sons, New York, 1977), which were then corrected for the engineered fusion and transit peptide to estimate the molecular weight of the corresponding native form.

Example 3

Monoterpene Synthase Assays and Product Identification

Monoterpene Synthase Assays. [1-3H]Geranyl diphosphate (250 Ci/mol) was prepared by an established method (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994). Terpenoid standards were from our own collection. Unless otherwise stated, all reagents were obtained from Sigma Chemical Co. or Aldrich Chemical Co.

Monoterpene synthase activities were assayed by methods previously described (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994; Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984; Croteau, R., and Cane, D.E., Methods Enzymol. 110:383-405, 1985). Briefly, an aliquot of the bacterial cell lysate,

appropriate column fractions, or partially purified and desalted enzyme preparation, in 0.5 or 1.0 ml of 50 mM Mopso buffer (pH 6.1 to 7.0, as appropriate for the target activity) containing 10 mM MgCl₂, 5 mM DTT and 10% (v/v) glycerol, was transferred to a 7 ml glass, Teflon sealed, screw-capped tube, and the mixture was overlayed with 1 ml pentane to trap volatile products. The reaction was initiated by the addition of 4.5 μM [1-3H]geranyl diphosphate (1.3 μCi), with incubation at 31°C with gentle shaking for 0.5 to 3.0 h. The pentane layer and an additional pentane extract (2 x 1 ml) were passed over a short column of silica gel surmounted by anhydrous MgSO₄ (in a Pasteur pipette) to afford the monoterpene olefin fraction. Subsequent extraction of the remaining aqueous phase with diethyl ether (2 x 1 ml), and passage of this extract through the same column, yielded the oxygenated monoterpene fraction. The residual aqueous phase was then treated with excess potato apyrase and wheat germ acid phosphatase to hydrolyze monoterpenol diphosphate esters (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Croteau, R., and Cane, D.E., Methods Enzymol. 110:383-405, 1985). The liberated alcohols were then extracted into diethyl ether (2 x 1 ml) and the combined extract dried over anhydrous MgSO₄. Radioactivity in the various fractions was determined by liquid scintillation counting of aliquots (Packard 460 CD with external standard quench correction) and the remaining material was concentrated for radio-GC and GC-MS analysis.

5

10

15

20

25

30

Kinetic analyses were carried out with the partially purified, recombinant pGEX fusion proteins by determination of initial reaction rates at a minimum of ten substrate concentrations ranging from 0.45 to 45 μM [1-3H]geranyl diphosphate, at saturating levels of the divalent metal ion cofactor. The results were analyzed by nonlinear regression of the Michaelis-Menten equation using the curve-fitting capabilities of Sigma-Plot (Jandel Corp.).

Product Identification - To obtain sufficient product for analysis by radio-GC and chiral phase capillary GC-MS, the samples from multiple assays were pooled as necessary. Radio-GC was performed on a Gow-Mac 550P gas chromatograph with thermal conductivity detector directly coupled to a Packard 894 gas proportional counter (30). An AT-1000 packed column (Alltech) was used with He as carrier at 30 ml/min and with temperature programming from 70°C to 200°C (at 5°C/min) for analysis of monoterpene olefins, and from 100°C to 200°C (at 5°C/min) for analysis of oxygenated monoterpenes. Authentic standards (10-20 μg/component) were

-31-

included with each injection in order to correlate retention times determined by mass and radioactivity detectors.

GC-MS was performed on a Hewlett-Packard 6890 GC-quadrupole mass selective detector system interfaced with a Hewlett-Packard Chemstation for data analysis. Chiral phase separations were performed by split injection (25:1) on a 30 m cyclodex-B capillary column (J&W Scientific) using He as carrier at 0.6 ml/min and temperature programming from 35°C to 200°C at 10°C/min.

5

10

15

20

25

30

35

Example 4

Product Profiles of Recombinant Synthases

Since the formation of multiple products from geranyl diphosphate is a common, if unusual, feature of the monoterpene synthases (Croteau, R., Chem. Rev. 87:929-954, 1987; Wise, M. L., and Croteau, R., in Comprehensive Natural Products Chemistry: Isoprenoids (Cane, D. E., ed) Vol. 2 (in press), Elsevier Science, Oxford, 1998), the product profiles of the recombinant enzymes were examined in detail by radio-GC and GC-MS. Recombinant sabinene synthase (SSS) produces exclusively monoterpene olefins, which by radio-GC were identified as sabinene (62%), K-terpinene (21%), terpinolene (6.7%), limonene (6.5%) and myrcene (2.5%). The major products of this enzyme (sabinene and K-terpinene) are formed by a cyclization mechanism involving a 1,2-hydride shift in the α-terpinyl cation intermediate. Chiral phase capillary GC-MS demonstrated the biosynthetic sabinene to be coincident with authentic (+)-sabinene; however, the (-)-enantiomer was not available for analysis to confirm the absolute configuration of this product. Previous studies have shown that cell-free extracts from sage produce only the (+)-antipode of sabinene from geranyl diphosphate (Croteau, R., in Recent Developments in Flavor and Fragrance Chemistry (Hopp, R., and Mori, K., eds), pp. 263-273, VCH, Weinheim, Germany, 1992; Croteau, R., in Flavor Precursors: Thermal and Enzymatic Conversions (Teranishi, R., Takeoka, G. R., and Guntert, M., eds), American Chemical Society Symposium Series, No. 490, pp. 8-20, Washington, DC, 1992), supporting the assignment of the (+)-stereoisomer in this case. The other principal olefinic products of SSS are achiral.

Cineole synthase (SCS) was shown by aliquot counting and radio-GC of the various metabolite fractions to produce both oxygenated monoterpenes (1,8-cineole, 79%, with a few percent α -terpineol) and a mixture of olefins (~20%). Chiral phase capillary GC-MS allowed resolution, confirmation and quantification of the olefins as (+)- α -pinene (5.5% of total products), (-)- α -pinene (0.9%), myrcene (2.9%),

sabinene (2.6%, presumably the (+)-enantiomer), (+)- β -pinene (2.7%), (-)- β -pinene (4.1%), (+)-limonene (1.1%) and (-)-limonene (0.4%). The stereochemistry of the enzymatic transformation leading to 1,8-cineole has been examined (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994) and shown to involve the cyclization of the bound intermediate 3R-linally diphosphate in anti,endo-conformation, i.e., the same overall stereochemistry required for the production of (+)- α -pinene, (+)- β -pinene and (+)-limonene (Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984). The formation of the (-)-series of antipodes must therefore occur via the extended (anti,exo) conformation. This apparent loss of stereochemical fidelity in the production of some of the olefin by-products may be a consequence of the fact that the enzyme is expressed as the pGEX fusion of the preprotein of the native synthase, and thus bears a large amino-terninal extension that could compromise substrate and intermediate binding conformations.

5

10

15

20

25

30

35

Bornyl diphosphate synthase (SBS) was shown, by radio-GC evaluation of all metabolite fractions, to produce principally bornyl diphosphate (75%), as demonstrated by enzymatic hydrolysis of this product followed by separation of the derived borneol from the residual geraniol (liberated from the substrate) and from lesser amounts of non-enzymatic solvolysis products (also generated from geranyl diphosphate in the course of the analysis). The production of bornyl diphosphate by this recombinant enzyme was also demonstrated directly by radio-HPLC analysis of the aqueous reaction mixture using an ion-paring, reversed-phase chromatography protocol previously established for the separation of prenyl diphosphate esters (McCaskill, D., and Croteau, R., Anal. Biochem. 215:142-149, 1993). Additionally, chiral phase capillary GC-MS analysis of the derived borneol demonstrated the exclusive presence of the (+)-antipode, as expected based on studies with the corresponding native enzyme (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Croteau et al., J. Biol. Chem. 260:5956-5962, 1985; Croteau et al., J. Biol. Chem. 261:13438-13445, 1986).

The recombinant (+)-bornyl diphosphate synthase was also shown, by radio-GC of the olefin fraction and chiral phase GC analysis, to produce a series of olefins (25% of total product) identified as (+)- α -pinene (3.4% of total product), (+)-camphene (9.5%), (-)-camphene (0.5%), (+)-limonene (3.9%), (-)-limonene (3.9%), terpinolene (2.1%) and myrcene (1.5%). Since formation of the (+)-olefin series is mechanistically related to the formation of (+)-bornyl diphosphate via the anti, endo-cyclization of the intermediate 3R-linalyl diphosphate (Croteau et al., J.

-33-

Biol. Chem. 264:2075-2080, 1989; Croteau et al., J. Biol. Chem. 260:5956-5962, 1985; Croteau et al., J. Biol. Chem. 261:13438-13445, 1986; Croteau et al., J. Biol. Chem. 263:10063-10071, 1988; Croteau et al., Arch. Biochem. Biophys. 277:374-381, 1990), the generation of small amounts of the antipodal (-)-camphene and (-)-limonene by the recombinant cyclase again suggests some loss of stereochemical fidelity in the overall reaction sequence.

10

15

20

25

30

35

Example 5

Sequence Analysis of Recombinant Synthases

DNA sequences were assembled and analyzed using GCG software (Wisconsin Package version 9.0, Genetics Computer Group (GCG), Madison, WI). Alignment of the deduced amino acid sequences of SBS clone 3C6 (SEQ ID No:1) (2025 bps, with an ORF of 1794 nts encoding 598 amino acids for a protein of 69.3 kDa and calculated pl of 6.06), SCS clone 3B5 (SEQ ID No:3) (1968 bps, with an ORF of 1773 nts encoding 591 amino acids for a protein of 69.4 kDa and calculated pI of 5.79), and SSS clone 3F25 (SEQ ID No:5) (1911 bps, with an ORF of 1767 nts encoding 589 amino acids for a protein of 68.9 kDa and calculated pI of 5.22), with the published sequences for (-)-limonene synthase from Mentha spicata (spearmint) (Colby et al., J. Biol. Chem. 268:23016-23024, 1993) and Perilla frutescens (Yuba et al., Arch. Biochem. Biophys. 332:280-287, 1996), linalool synthase from Clarkia breweri (Dudareva et al., Plant Cell 8:1137-1148, 1996), and three monoterpene olefin synthases from Abies grandis (grand fir) (Bohlmann et al., J. Biol. Chem. 272:21784-21792, 1997), illustrates that there are several regions of similarity between these nine monoterpene synthases of diverse origin. Comparison of these sequences using the GCG GAP program (Wisconsin Package version 9.0, Genetics Computer Group (GCG), Madison, WI) revealed the monoterpene synthases from sage to resemble each other and the limonene synthases from related members of the Lamiaceae (50-70% identity, 70-85% similarity) more closely than the monoterpene synthases of the gymnosperm grand fir (-32% identity) or the linalool synthase from C. breweri (-25% identity).

Monoterpene biosynthesis is compartmentalized in plastids (Gleizes et al., Planta 159:373-381, 1983; Mettal et al., Eur. J. Biochem. 170:613-616, 1988; Perez et al., Plant Physiol. Biochem. 28:221-229, 1990), thus the monoterpene synthases are encoded as preproteins bearing an amino-terminal transit peptide for import of these nuclear gene products into plastids (leucoplasts of the oil gland cells in the present instance) where they are proteolytically processed to the mature forms

(Keegstra et al., Annu. Rev. Plant Physiol. Plant Mol. Biol. 40:471-501, 1989). In all of the monoterpene synthases thus far examined, the 50 to 60 amino terminal residues are characterized by a low degree of similarity, typical of targeting sequences, yet they all share common features of transit peptides in being rich in serine, threonine and small hydrophobic residues but with few acidic residues (Keegstra et al., Annu. Rev. Plant Physiol. Plant Mol. Biol. 40:471-501, 1989; von Heijne et al., Eur. J. Biochem. 180:535-545, 1989).

5

10

15

20

25

30

35

All native monoterpene synthases thus far examined appear to be N-terminally blocked, preventing direct determination (by sequencing) of the transit peptide-mature protein cleavage junction (Lewinsohn et al., Arch. Biochem. Biophys. 293:167-173, 1992; McGeady, P., and Croteau, R., Arch. Biochem. Biophys. 317:149-155, 1995; Steele et al., Proc. Natl. Acad. Sci. USA 92:4164-4168, 1995; Colby et al., J. Biol. Chem. 268:23016-23024, 1993). Significantly, a tandem pair of arginine residues (e.g., arg⁵⁵ arg⁵⁶ of SBS, arg⁵⁷ arg⁵⁸ of SCS and arg⁵² arg⁵³ of SSS) are strictly conserved in the deduced sequences of all of the monoterpene synthases and they define the most N-terminal region of obvious homology, suggesting a possible cleavage site. It is believed that truncation of the recombinant (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase preproteins immediately upstream of these tandem arginines will yield fully functional "pseudomature" forms of the enzymes, whereas truncation downstream from this element will severely impair activity.

Downstream of the aforementioned, tandem arginines are several regions of homology, including the universally conserved (I,L or V)DDXXD (SEQ ID No:13) motif (e.g., residues I³⁵⁰-D³⁵⁵ of SBS) (SEQ ID No:2) found in virtually all deduced sequences for enzymes that utilize prenyl diphosphate substrates (Chen et al., *Protein Sci.* 3:600-607, 1994; Chen et al., *Arch. Biochem. Biophys.* 324:255-266, 1995). This aspartate rich element is now generally recognized as a binding site for the metal ion chelated diphosphate ester substrate (Chen et al., *Protein Sci.* 3:600-607, 1994; Ashby, M. N., and Edwards, P. A., *J. Biol. Chem.* 265:13157-13164, 1990; Tarshis et al., *Biochemistry* 33:10871-10877, 1994; Cane et al., *Biochemistry* 33:5846-5857, 1994; Tarshis et al., *Proc. Natl. Acad. Sci. USA* 93:15018-15023, 1996). Several other highly conserved regions are also apparent including, with reference to the amino acid sequence of SBS (SEQ ID No:2): Arg²⁹⁸-Trp-Trp³⁰⁰, Arg³⁷²-Trp-Glu/Gln³⁷⁴, Tyr³⁸⁴-Met-Gln/Lys³⁸⁶ and Cys⁵¹⁶-Tyr-Met-X-Glu/Asp⁵²⁰ (SEQ ID No:14). The active site peptide LQLYEASFLL (SEQ ID No:15), previously isolated

5

10

15

20

25

30

35

from the co-purified (+)-pinene synthase and (+)-bornyl diphosphate synthase of sage (McGeady, P., and Croteau, R., Arch. Biochem. Biophys. 317:149-155, 1995) was located at residues 195-204 of SBS (SEQ ID No:2) and also at residues 187-196 of SSS (SEQ ID No:6). Very similar sequences in the same location were found in SCS (SEQ ID No:4, amino acid residues 191-200) and in the two limonene synthase sequences from M. spicata and P. frutescens.

The (+)-bornyl diphosphate synthase from sage has previously been shown to be inhibited by the 'active serine'-directed reagent diisopropylfluorophosphate (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979), a characteristic not shared by other monoterpene cyclases (Alonso, W. R., and Croteau, R., in Methods in Plant Biochemistry (Enzymes of Secondary Metabolism) (Lea, P.J., ed) Vol. 9, pp. 239-260, Academic Press, New York, 1993). Because of the unique utilization of the substrate diphosphate moiety as the terminating nucleophile by this enzyme (Cane et al., J. Am. Chem. Soc. 104:5831-5833, 1982, Croteau et al., Biochemistry 24:7077-7085, 1985), it was hypothesized that a serine residue may be involved in binding and transfer of the diphosphate function in the Sequence comparison of SBS with the other two course of the reaction. monoterpene synthases of sage reveals four unique serine residues at positions 302, 320, 454 and 469 (SEQ ID No:2). Two (at positions 302 and 320) are within otherwise highly conserved regions and are, therefore, obvious targets for selective covalent modification with radiolabeled diisopropylfluorophosphate and directed mutagenesis studies.

Example 6

Physical Properties of Recombinant Synthases

Properties of SBS Calibrated gel permeation chromatography of the pGEX fusion form of SBS revealed a single peak of activity at an elution volume corresponding to an M_r -200,000, indicating that the expressed fusion preprotein (corresponding to a molecular weight of about 2 x 96,300) was a functional dimer. Treatment of the SBS protein with thrombin to remove the glutathione-S-transferase fusion tag, followed by re-chromatography, indicated a decrease in molecular weight to approximately 135,000, consistent with the loss of the 27 kDa transferase peptide from each subunit at a calculated molecular weight of 69,300 for the preprotein. Further correction of the molecular weight to account for the transit peptide would yield a dimer of about 120 kDa which corresponds roughly to the native dimer molecular weight of both (+)-bornyl diphosphate synthase and (+)-pinene synthase

from sage (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984), two enzymes which have never been satisfactorily resolved as distinct species. Although a dimeric quarternary structure is not unique to these two synthases, the vast majority of the monoterpene synthases characterized to date are monomeric (Alonso, W. R., and Croteau, R., in Methods in Plant Biochemistry (Enzymes of Secondary Metabolism) (Lea, P. J., ed) Vol. 9, pp. 239-260, Academic Press, New York, 1993).

5

10

15

20

25

30

35

The product profile of the protein encoded by SBS clone 3C6 (SEQ ID No:2) is qualitatively similar to the combination of both (+)-bornyl diphosphate synthase and (+)-pinene synthase (i.e., (+)-bornyl diphosphate and the (+)-series of α-pinene and related olefins) (Croteau, R., and Karp, F., Arch. Biochem. Biophys. 198:512-522, 1979; Gambliel, H., and Croteau, R., J. Biol. Chem. 259:740-748, 1984), although the quantitative distributions do not exactly match, and the stereochemistry of the olefin products is anomalous. Thus, (+)-bornyl diphosphate and (+)-α-pinene, (+)-camphene and (+)-limonene arise via the same overall cyclization stereochemistry, and these enantiomers are produced exclusively from geranyl diphosphate by the native (+)-bornyl diphosphate and (+)-pinene synthase activities (Croteau et al., J. Biol. Chem. 264:2075-2080, 1989; Croteau et al., J. Biol. Chem. 260:5956-5962, 1985; Croteau et al., J. Biol. Chem. 261:13438-13445, 1986; Croteau et al., J. Biol. Chem. 263:10063-10071, 1988; Croteau et al., Arch. Biochem. Biophys. 277:374-381, 1990).

The small amounts of (-)-limonene and (-)-camphene formed by the recombinant enzyme are attributed to antipodal cyclizations via abnormal extended conformations, as the phenomenon has been described previously, especially when using neryl diphosphate (the cis-analog of geranyl diphosphate) as an alternate substrate (Croteau et al., J. Biol. Chem. 263:10063-10071, 1988; Croteau, R., and Satterwhite, D. M., J. Biol. Chem. 264:15309-15315, 1989). The geranyl substrate, however, was verified as >99% pure, thereby eliminating this possibility in the present instance and suggesting that loss of stereochemical fidelity (to the extent of 5% of the total product mixture) may be attributed to the presence of the glutathione-S-transferase fusion peptide plus transit peptide which may alter substrate binding directly, or indirectly by compromising subunit assembly.

Thus, the physical properties of the recombinant (+)-bornyl diphosphate synthase, together with the distribution and stereochemistry of its products, suggest that this enzyme might represent both (+)-bornyl diphosphate synthase and (+)-pinene

synthase which were previously assumed to be distinct enzymes. The resolution of this question will require the detailed assessment of truncated enzymes that more closely resemble the native form, and which will therefore be likely to produce the same mixture of monoterpenes as the native form.

5

10

15

20

25

30

35

Physical Properties of SCS and SSS. Gel permeation chromatography of SCS revealed a single peak of activity at an elution volume corresponding to an M_r of 72,000, whereas SSS gave two peaks of activity, an aggregated form eluting in the void volume and a second corresponding to an M_r of 60,000. Both of these molecular weights are significantly lower than those predicted from pGEX expression-based fusion of the glutathione-S-transferase (27 kDa) with the respective preproteins (SCS ~96 kDa and SSS ~96 kDa). Thrombin treatment was without influence on the gel permeation chromatographic behavior of these enzymes, indicating the absence of the glutathione-S-transerase peptide tag and rationalizing the previously observed inability of the recombinant SCS and SSS enzymes to bind to the glutathione affinity column. Inspection of the 5'-sequences of the corresponding pGEX constructs showed both to be free of in-frame stop codons that might have permitted polycistronic translation of the preprotein devoid of the glutathione-Stransferase peptide. The apparent truncation was therefore attributed to proteolytic processing of the recombinant SCS and SSS in the E. coli host to proteins that seemingly resemble the preprotein forms of the native, monomeric, sage 1,8-cineole synthase (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994; Croteau, R., and Karp, F., Arch. Biochem. Biophys. 179:257-265, 1977) and (+)-sabinene synthase (Croteau, R., in Recent Developments in Flavor and Fragrance Chemistry (Hopp, R., and Mori, K., eds), pp. 263-273, VCH, Weinheim, Germany, 1992; Croteau, R., in Flavor Precursors: Thermal and Enzymatic Conversions (Teranishi, R., Takeoka, G.R., and Guntert, M., eds), American Chemical Society Symposium Series, No. 490, pp. 8-20, Washington, DC, 1992). Similar proteolytic processing of a recombinant limonene synthase preprotein from spearmint has been observed previously in this E. coli host (Colby et al., J. Biol. Chem. 268:23016-23024, 1993).

1,8-Cineole synthase has never been satisfactorily separated from the aforementioned (-)-pinene synthase from sage but, in this instance, the product distribution of SCS does not match well the product distribution of (-)-pinene synthase either quantitatively, qualitatively, or in stereochemical terms, since the reactions catalyzed are of the opposite antipodal series (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994; Gambliel, H., and Croteau, R., J. Biol. Chem.

257:2335-2342, 1982; Croteau et al., J. Biol. Chem. 264:2075-2080, 1989). However, the product distribution of SCS shows some parallels with that of the recently described cyclase III which produces (+)-α-pinene and (+)-β-pinene (Wagschal et al., Arch. Biochem. Biophys. 308:477-487, 1994; Pyun et al., Arch. Biochem. Biophys. 308:488-496, 1994). Even here, the match is not perfect and the production of anomalous products of the antipodal (-)-series (<6% of total) again suggests that substrate binding interactions may be compromised by the presence of the substantial transit peptide.

To assess the latter possibility, the K_m values for SCS (7.0 μ M), SSS (7.4 μ M) and SBS (3.0 μ M) were determined. These values are likely somewhat high because the recombinant enzymes were not purified sufficiently to remove all contaminating phosphatases that result in some depletion of the substrate geranyl diphosphate. While the calculated K_m values compare reasonably well with the literature values of 1.1 μ M (Croteau et al., Arch. Biochem. Biophys. 309:184-192, 1994), 2.0 μ M (Croteau, R., in Recent Developments in Flavor and Fragrance Chemistry (Hopp, R., and Mori, K., eds), pp. 263-273, VCH, Weinheim, Germany, 1992) and 2.0 μ M (Croteau et al., J. Biol. Chem. 264:2075-2080, 1989; Croteau et al., Arch. Biochem. Biophys. 277:374-381, 1990), respectively, for the corresponding native enzymes, they are sufficiently higher to suggest at least subtle alteration in binding capacity of the recombinant forms.

10

15

20

25

30

35

Example 7

Hybridization Conditions and Representative Sequences

The nucleic acid sequences of the present invention that encode at least a portion of a (+)-bornyl diphosphate synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:1, are capable of hybridizing to the nucleic acid sequence set forth in SEQ ID NO:1, or to its complementary sequence, under the following conditions: 3x SSC at 65°C for 16 hours.

The nucleic acid sequences of the present invention that encode at least a portion of a (+)-bornyl diphosphate synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:1, are capable of remaining hybridized to the nucleic acid sequence set forth in SEQ ID NO:1, or to its complementary sequence, under the following conditions: two washes (twenty minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C; preferably two washes (twenty minutes per wash) in 2x SSC at a temperature in the

range of from 18°C to 24°C, followed by one wash in 0.5x SSC for 30 minutes at 55°C; most preferably two washes (fifteen minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C, followed by two washes (twenty minutes per wash) in 0.2x SSC at a temperature of 65°C.

5

10

15

20

25

30

35

The nucleic acid sequences of the present invention that encode at least a portion of a (+)-sabinene synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:5, are capable of hybridizing to the nucleic acid sequence set forth in SEQ ID NO:5, or to its complementary sequence, under the following conditions: 3x SSC at 65°C for 16 hours.

The nucleic acid sequences of the present invention that encode at least a portion of a (+)-sabinene synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:5, or to its complementary sequence, are capable of remaining hybridized to the nucleic acid sequence set forth in SEQ ID NO:5, or to its complementary sequence, under the following conditions: two washes (twenty minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C; preferably two washes (twenty minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C, followed by one wash in 0.5x SSC for 30 minutes at 55°C; most preferably two washes (fifteen minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C, followed by two washes (twenty minutes per wash) in 0.2x SSC at a temperature of 65°C.

The nucleic acid sequences of the present invention that encode at least a portion of a 1,8-cineole synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:3, are capable of hybridizing to the nucleic acid sequence set forth in SEQ ID NO:3, or to its complementary sequence, under the following conditions: 3x SSC at 65°C for 16 hours.

The nucleic acid sequences of the present invention that encode at least a portion of a 1,8-cineole synthase protein, or that are complementary to at least a portion of the nucleic acid sequence set forth in SEQ ID NO:3, or to its complementary sequence, are capable of remaining hybridized to the nucleic acid sequence set forth in SEQ ID NO:3, or to its complementary sequence, under the following conditions: two washes (twenty minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C; preferably two washes (twenty

5

10

15

20

25

30

35

minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C, followed by one wash in 0.5x SSC for 30 minutes at 55°C; most preferably two washes (fifteen minutes per wash) in 2x SSC at a temperature in the range of from 18°C to 24°C, followed by two washes (twenty minutes per wash) in 0.2x SSC at a temperature of 65°C.

Representative nucleic acid sequences of the present invention that encode (+)-bornyl diphosphate synthase are set forth in SEQ ID NO:1, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:30, SEQ ID NO:32. Each of the foregoing, representative nucleic acid sequences encode the same (+)-bornyl diphosphate synthase protein sequence.

Representative (+)-bornyl diphosphate synthase proteins of the present invention are set forth in SEQ ID NO:2, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41 and SEQ ID NO:42.

Representative nucleic acid sequences of the present invention that encode (+)-sabinene synthase are set forth in SEQ ID NO:5, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59. Each of the foregoing, representative nucleic acid sequences encode the same (+)-sabinene synthase protein sequence.

Representative (+)-sabinene synthase proteins of the present invention are set forth in SEQ ID NO:6, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69.

Representative nucleic acid sequences of the present invention that encode 1,8-cineole synthase are set forth in SEQ ID NO:3, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84 and SEQ ID NO:86. Each of the foregoing, representative nucleic acid sequences encode the same 1,8 cineole synthase protein sequence.

Representative 1,8-cineole synthase proteins of the present invention are set forth in SEQ ID NO:4, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID NO:96 and SEQ ID NO:97.

It will be recognized that, based on the disclosure of the present application, one of ordinary skill in the art can readily identify and generate numerous amino acid

-41-

sequences and nucleic acid sequences that fall within the scope of the present invention. Thus, by way of non-limiting example, one of ordinary skill in the art will recognize that it is possible to make numerous, conservative amino acid substitutions in the protein sequences set forth, for example, in SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:6 which do not significantly adversely affect the enzymatic activity of the modified proteins. The term "conservative substitution" refers to substituting an amino acid with a side chain that is similar in charge and/or structure to that of the native molecule.

5

10

15

Further, one of ordinary skill in the art will recognize that, with the aid of a genetic code table, it is possible to generate numerous nucleic acid sequences that encode any or all of the (+)-bornyl diphosphate synthase, 1,8-cineole synthase and (+)-sabinene synthase protein sequences disclosed herein.

While the preferred embodiment of the invention has been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

- 1. An isolated nucleic acid sequence encoding (+)-bornyl diphosphate synthase, 1,8-cineole synthase or (+)-sabinene synthase from an angiosperm plant species.
- 2. An isolated nucleic acid sequence of Claim 1 encoding (+)-bornyl diphosphate synthase.
- 3. An isolated nucleic acid sequence of Claim 2 encoding (+)-bornyl diphosphate synthase from an essential oil plant species.
- 4. An isolated nucleic acid sequence of Claim 3 encoding (+)-bornyl diphosphate synthase from Salvia officinalis.
- 5. An isolated nucleic acid sequence of Claim 4 consisting of the nucleic acid sequence set forth in SEQ ID NO:1.
- 6. An isolated nucleic acid sequence of Claim 1 encoding 1,8-cineole synthase.
- 7. An isolated nucleic acid sequence of Claim 6 encoding 1,8-cineole synthase from an essential oil plant species.
- 8. An isolated nucleic acid sequence of Claim 7 encoding 1,8-cineole synthase from Salvia officinalis.
- 9. An isolated nucleic acid sequence of Claim 8 consisting of the nucleic acid sequence set forth in SEQ ID NO:3.
- 10. An isolated nucleic acid sequence of Claim 1 encoding (+)-sabinene synthase.
- 11. An isolated nucleic acid sequence of Claim 10 encoding (+)-sabinene synthase from an essential oil plant species.

- An isolated nucleic acid sequence of Claim 11 encoding (+)-sabinene 12. synthase from Salvia officinalis.
- An isolated nucleic acid sequence of Claim 12 consisting of the nucleic 13. acid sequence set forth in SEQ ID NO:5.
- A replicable expression vector comprising a nucleic acid sequence 14. encoding (+)-bornyl diphosphate synthase, 1,8-cineole synthase or (+)-sabinene synthase from an angiosperm plant species.
- A replicable expression vector of Claim 14 wherein the nucleotide 15. sequence comprises the sequence of SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5.
 - A host cell comprising a vector of Claim 14 or Claim 15. 16.
- A method of altering the expression of (+)-bornyl diphosphate 17. synthase in a suitable host cell comprising introducing into the host cell an expression vector that comprises a nucleotide sequence encoding an angiosperm (+)-bornyl diphosphate synthase protein under conditions enabling expression of the protein in the host cell.
- A method of altering the expression of 1,8-cineole synthase in a 18. suitable host cell comprising introducing into the host cell an expression vector that comprises a nucleotide sequence encoding an angiosperm 1,8-cineole synthase protein under conditions enabling expression of the protein in the host cell.
- A method of altering the expression of (+)-sabinene synthase in a 19. suitable host cell comprising introducing into the host cell an expression vector that comprises a nucleotide sequence encoding an angiosperm (+)-sabinene synthase protein under conditions enabling expression of the protein in the host cell.
- An isolated nucleic acid sequence that hybridizes, in a solution 20. comprising 3x SSC at a temperature of 65°C, to any one of the nucleic acid sequences set forth in SEQ ID NO:1, SEQ ID NO:3 and SEQ ID NO:5, said isolated nucleic acid sequence remaining hybridized to any one of the nucleic acid sequences set forth in SEQ ID NO:1, SEQ ID NO:3 and SEQ ID NO:5 in a solution comprising 0.5x SSC at a temperature of 55°C.

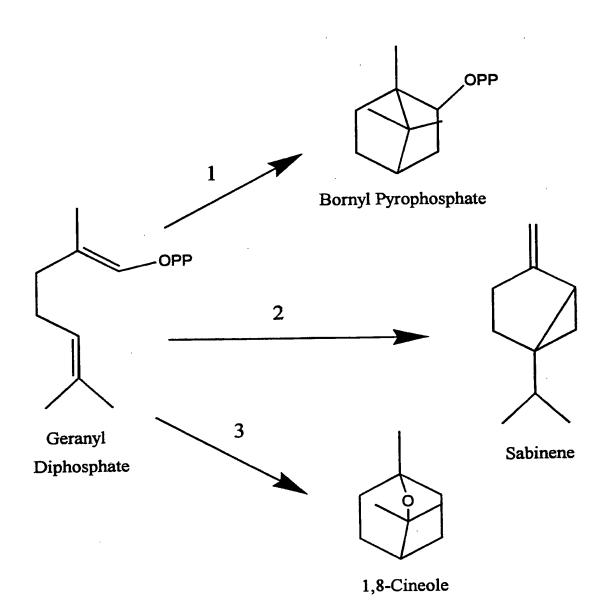


FIGURE 1

5

SEQUENCE LISTING

<110> Croteau, Rodney B Wise, Mitchell L Savage, Thomas J Katahira, Eva J <120> Monoterpene Synthases from Common Sage (Salvia officinalis) <130> wsur12893 <140> <141> <150> 08/937,540 <151> 1997-09-25 <160> 97 <170> PatentIn Ver. 2.0 <210> 1 <211> 2025 <212> DNA <213> Salvia officinalis <221> CDS <222> (11)..(1804) <400> 1 gatcacaaaa atg tot atc att agc atg aac gta tog atc ott agc aag Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys CCa cta aat tgc ctc cac aac ttg gag agg aga cct tca aaa gcc ttg Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu 20 ett gte eet tge act gea eee ace get ege ete egg gea tet tge tee 145 . Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser tca aaa cta caa gaa gct cat caa atc cga cga tct gga aac tac caa 193 Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln cct gce ctt tgg gat tcc aat tac att cag tct ctc aat act cca tat 241 Pro Ala Leu Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr 70 acg gag gag agg cac ttg gat aga aaa gca gag ctg att gtg caa gtg 289 Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val agg ata ctg cta aag gaa aaa atg gag cct gtt caa caa ttg gag ttg Arg Ile Leu Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu 337 95 100

at: 110		t ga s As	c tt p Le	g aa u Ly	a ta s Ty 11	r Lei	1 Gly	g cto y Let	tco 1 Se	g gar As _l 120	P Pho	t tti	t ca a Gl	a ga n As _l	t gag p Glu 125	385
at:	t aag e Ly:	g ga s Gl	g at u Il	c tt e Le 13	n eti	t gtt y Val	ata Ile	tac Tyr	aat Asr 135	ı Glı	g cad 1 His	aaa Lys	tge Cy:	s Phe	t cac e His	433
aat Asr	aat 1 Asi	ga Gl:	a gta u Val 14!	r GT	g aaa u Lys	atg Met	gat Asp	tto Leu 150	. Tyr	tto Phe	aca Thi	gct Ala	cti Lei 155	i Gly	t ttc y Phe	481
a ga Arg	cto Lev	te Lev 160	a wri	a caa g Glr	a cat n His	ggt Gly	ttt Phe 165	Asn	att Ile	Ser	caa Gln	gat Asp 170	Val	ttt Phe	aat Asn	529
tgt Cys	Phe 175	пy	g aac 3 Asn	gaç Glu	g aag 1 Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	agc Ser	Leu	gct Ala	Caa Gln	57 7
gat Asp 190	****	aac Lys	gga Gly	atg Met	tta Leu 195	caa Gln	ctg Leu	tat Tyr	gaa Glu	gcg Ala 200	Ser	ttc Phe	ctt Leu	ttg Leu	aga Arg 205	625
aaa Lys	ggt Gly	gaa Glu	gat Asp	Thr 210	Leu	gag Glu	ctt Leu	gca Ala	aga Arg 215	gaa Glu	ttt Phe	gcc Ala	aca Thr	aaa Lys 220	tgt Cys	673
ctg Leu	cag Gln	aaa Lys	aaa Lys 225	reu	gat Asp	gaa Glu	ggt Gly	ggt Gly 230	aat Asn	gaa Glu	att Ile	gat Asp	gag Glu 235	aat Asn	cta Leu	721
nec	Leu	240	lle	Arg	His	tct Ser	Leu 245	Asp	Leu	Pro	Leu	His 250	Trp	Arg	Ile	769
JIII	255	vai	GIU	Ala	Arg	tgg Trp 260	Phe	Ile	Asp	Ala	Tyr 265	Ala	Arg	Arg	Pro	817
270	Mec	ASII	PIO	Leu	275	ttc Phe	Glu	Leu	Ala	Lys 280	Leu	neA	Phe	Asn	11e 285	865
	5211	~	1111	290	GIN	caa Gln	GIU	Leu	Lys 295	Asp	Leu	Ser	Arg	Trp 300	Trp	913
		200	305	rne	PIO	gaa Glu	rys	310	Pro	Phe	Val .	Arg .	Asp 315	Arg	Leu	961
		320		- 116	11p		325	GTA 1	Met .	rne	Glu	330	His	Gln	His	1009
3	tat Tyr 335	cag Gln	aga Arg	aaa Lys	met.	gcc (Ala 2 340	gcc a Ala 1	aca a Thr	att a Ile :	Ile '	gtt : Val : 345	tta (Leu)	gca Ala	aca Thr	gtt Val	1057

at: 11: 35:		t ga p As	at at sp Il	t ta Le Ty	ic ga 'r As 35	p va.	g tat L Tyr	gg Gl	t ac y Th	a ct r Le 36	u As	t gaa p Gl	a ct. u Le	a gaa u Gli	a cta u Leu 365	1105
tt: Pho	ac.	a ga r As	c ac	g tt r Ph 37	е гу	g aga s Arg	tgg Trp	gai As	t ac P Th 37	r GI	a tc u Se:	a ata r Ile	a aco	c cga Arg 380	, Leu	1153
Pro	tat Ty	t ta Ty	c at r Me 38	C GI	a tt. n Le	a tgt u Cys	tat Tyr	tgg Tr <u>r</u> 390	o GT	t gte y Val	c cad	aac Asn	tat Ty:	: Ile	tcc Ser	1201
٠		40	0	LAS	5 116	t ctc = Leu	405	GIT	ı His	s Gl	/ Phe	Phe 410	Cys	Leu	Gln	1249
tat Tyr	Cto Leu 415		g aa g Ly:	a tog s Sei	g gto CVal	g gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	Tyr	ttt Phe	cac His	gag Glu	1297
gca Ala 430	aag Lys	tg:	y tao Ty:	cac His	ago Ser 435	ggt	tat Tyr	aca Thr	CCa Pro	ago Ser 440	Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
aac Asn	atc Ile	Ala	aaq Lys	I att	ser	gtg Val	gcg Ala	tct Ser	cct Pro 455	gca Ala	ata Ile	ata Ile	tcc Ser	cca Pro 460	acc Thr	1393
-3-		****	465	. Ala	Asn	gcg Ala	Ser	470	Asp	Thr	Ala	Val	11e 475	qzA	Ser	1441
	-3-	480	TYL	urs	Asp	ata Ile	485	Cys	Leu	Ala	Gly	11e 490	Ile	Leu	Arg	1489
	495	· wp	nep	Deu	GIY	aca Thr 500	ser	Tyr	Pne	Glu	Leu 505	Ala	Arg	Gly	Asp	1537
gtg Val 510	ecg Pro	aaa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr 1	atg Met	aag Lys	gaa Glu 520	aca Thr	aat Asn	gct Ala	Ser	gag Glu 525	1585 [.]
			Val	530	urz	gtg Val	ràs i	Phe	Leu 535	Ile	Arg	Glu .	Ala	Trp 540	Lys	1633
gat a Asp 1	atg Met .	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca (Ala /	-Ta (ggt Sly 550	tat Tyr	ccg Pro	ttt Phe	Pro A	gat Asp 555	ggt Gly i	atg Met	1681
gtg (ggc Gly 560	gca Ala	gct Ala	aat Asn	TIE (ggg o Gly A	ege (gtg Val .	gcg Ala	GIn	ttt a Phe 1 570	itt [le '	tat (Tyr)	ctc Leu	1729
Cac o	ga Sly 1 175	gat Asp	ej A aaa	ttt Phe		gtg d Val d 580	aa c Sln H	ac (tcg (Ser)	rAa .	acg thr 5	tac ç Tyr (gag (Slu)	cat a	atc [le	1777

gcc ggc cta ctg ttc gag cct tat gca tgaacaaatg ggagactgct
Ala Gly Leu Leu Phe Glu Pro Tyr Ala
590 595

tgatatatat taatttggca caccaataat tgcatgttat atatgttgga aaataagtgt 1884
ctggttgaga tgtcatgtgg tgtattatct aaataattca aggttgcctt gtttatgtag 1944

agetegagga aaaaaaaaa a 2025

<210> 2 <211> 598 <212> PRT <213> Salvia officinalis

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155 160

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 215 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 230 235 240

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 290 295 300

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 315

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 340 345 350

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355 360 365

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 370 380

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg 405 410 415

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp
420 425 430

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp
485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu Leu Phe Glu Pro Tyr Ala 595 <210> 3 <211> 1968 <212> DNA <213> Salvia officinalis <220> <221> CDS <222> (14)..(1786) ggatcaccac aag atg tcg agt ctt ata atg caa gtt gtt att cct aag 49 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys cca gcc asa att ttt cac aat asc tta ttc agc gtg att tca asa cga Pro Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg cat cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat His Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His tgc tca cta caa atg ggt aat gag atc caa act gga cga cga act gga Cys Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly ggc tac cag cct acc ctt tgg gat ttc agc acc att caa ttg ttc gac Gly Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp tot gag tat aag gaa gag aag cac ttg atg agg goo goa ggt atg ata 289 Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile gcc caa gtg aat atg ttg ttg cag gaa gaa gta gat tcg att caa cgg 337 Ala Gln Val Asn Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg 100 ttg gag ttg att gat gac cta cga agg ctg ggt ata tct tgc cat ttt 385 Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe gac ege gag ate gtt gaa ata tta aac tea aaa tat tat ace aac aat 433 Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn 135

		- 1.01	9 61	14.	5 AS	Ъ те	ц ту	r Se.	150	r Ala	1 Lei	ı Arg	Phe	15:	_	481
cta Leu	aga Ar	a caa g Glr	tac Ty:	r wai	t tt	t ago e Sei	gto Val	tci Sei 165	C GTI	a gag n Glu	g gta 1 Val	ttt. Phe	gat Asp 170	Cys	t ttc 3 Phe	529
aag Lys	aat Asn	: gac Азр 175	, Lys	Gly Ggt	act	t gat r Asp	tto Phe 180	: Lys	g cca Pro	ago Ser	cta Leu	gtc Val 185	gat Asp	gat Asp	act Thr	577
3	190	Deu	. Deu	GII	Let	195	GIU	Ala	Ser	Phe	Leu 200	Ser	Ala	Gln	ggc Gly	625
205		****	Leu	nis	210	ALA	Arg	Asp	Phe	gct Ala 215	Thr	Lys	Phe	Leu	His 220	673
-4-		4 67	neu	225	Asp	Lys	Asp	Ile	Asn 230	ctc Leu	Leu	Ser	Ser	Ile 235	Glu	721
9		200	240	beu	PIO	The	HIS	245	Arg	gtt Val	Gln	Met	Pro 250	Asn	Ala	769
		255	116	ASP	ATA	Tyr	260	Arg	Arg	ccc Pro	Asp	Met 265	Asn	Pro	Thr	817
	270	J1 1	neu	Ala	Lys	275	Asp	rne	Asn	atg Met	Val 280	Gln .	Ala	Gln	Phe	865
Gln (285		G.L.	neu	Lys	290	ALA	ser	Arg	Trp	Trp 295	Asn	Ser '	Thr	Gly	Leu 300	913
gtc (Val :		J	Deu	305	rne	val	Arg .	Asp	Arg 310	Ile	Val (Glu (Cys '	Tyr 315	Tyr	961 ·
tgg a			320	Val	Val	GIU.	Arg A	325	Glu	His (GTA .	Tyr C	31u / 330	Arg	Ile	1009
atg c Met L		335	uys	116	ASN	Ala .	340	val .	Thr '	Thr :	Ile A	Asp A 345	(sp	/al	Phe	1057
	50		y .	• • • • •	ueu	355	ara 1	Leu (oin i	Leu I	?he 1 360	hr T	hr A	lla :	Ile	1105
caa a Gln A 365	ga t rg 1	gg g	jat a Asp]		gaa Glu B70	tca a Ser N	atg a Met I	ag d Lys (ern 1	tc d Leu E 375	ct c Pro P	ct t	ac a yr M	let (caa Gln 380	1153

ata Ile	a tg e Cy	t ta s Ty	t ctt r Leu	get 1 Ala 385	r rer	ttc Phe	aac Ası	ttt Phe	gte Val 390	L Asr	gag 1 Glu	g atg 1 Met	gct Ala	tat Tyr 395		1201
act Thi	Ct: Le	t agg	g gat g Asp 400	rys	ggt Gly	ttc Phe	aac Asn	Ser 405	Thr	Pro	tat Tyr	cta Leu	cga Arg 410	Lys	gcg	1249
tg <u>c</u> Trp	gti Val	gat L Asp 415	ttg Leu S	gtt Val	gag Glu	tca Ser	tat Tyr 420	Leu	ata Ile	gag Glu	gca Ala	aag Lys 425	tgg Trp	tac Tyr	tac Tyr	1297
atg Met	gga Gly 430		aaa Lys	Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atg Met	aag Lys 440	Asn	agt Ser	tgg Trp	ata Ile	1345
tca Ser 445		gga Gly	ggc,	atc Ile	ccc Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 455	ttt Phe	ttc Phe	cgg Arg	cta Leu	aca Thr 460	1393
gat Asp	tcg Ser	att	gag Glu	gaa Glu 465	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	1441
att Ile	gtt Val	cgt Arg	gca Ala 480	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	1489
	-	495	gat Asp	GIU	Val	GIU	Arg 500	СТĀ	Asp	Val	Pro	Lys 505	Ser	Val	Gln	1537
-,5	510	Mec	aat Asn	GIU	гÀг	515	Ala	Ser	Glu	Glu	61u 520	Ala .	Arg	Glu	His	1585
525	.ug	DGL	Ctc Leu	116	530	GIN	rnr	Trp	Lys	Met 535	Met	Asn	Lys	Glu	Met 540	1633
	****	Set		545	ser	ras :	ľyr	Phe	Val 550	Gln	Val	Ser 1	Ala .	Asn 555	Leu	1681
		••••	gcg Ala 560	G111	rp	TIE :	ryr	565	His	Glu	Ser .	Asp (31y 1 570	Phe (Gly	1729
atg Met		cat His 575	tca (Ser)	ttg (Leu 1	gtg : Val :	ASD I	iys i 80	atg (Met]	ctc . Leu .	aga (Arg (Gly 1	ttg t Leu I 585	tg (Leu I	ttc (Phe 1	gac Asp	1777
	tat Tyr 590	gag Glu	taact	caato	et t	cgcco	ggg	t tc	caaa	tgaa	tcaa	atctq	ŗtt			1826
															gccc	
aacc	gtca	gg c	ccatt	tcg	tc:	atgtt	cat	aata	aata	aat a	aaaa	etgtt	a at	caat	aaca	1946

aaaaaaaaa aaaaaaaaaa aa

1968

<210> 4 <211> 591 <212> PRT <213> Salvia officinalis

<400> 4

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile
1 5 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 60 .

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 245 250 255

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 260 265 270

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 490 Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gin Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570 575 Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 585

<2 <2 <2 <2	12> 13> 20>	1911 DNA Salv		ffic	inal	is										
<2	21> 22>	CDS (25)	(1	791)												
	00> aata	_	aac	taac	aat	aaaa	ato	tct	tee	att	age	ata	226	2+2	***	£ 1
							Met 1	Ser	Ser	Ile	Ser 5	Ile	Asn	Ile	Ala	51
10)	J DC	u nsi	ı sei	1	5	s Asr	1 Phe	: Gli	20 20	; Ly:	s Pro	Sez	r Lys	a gca 3 Ala 25	99
			- 561	30)	r Ala	. Pro	ALA	35	Arg	Leu	ı Arg	Ala	Ser 40		14
			45	GIU	. Lys	, PIC) HIS	50	ııe	Arg	Arg	, Ser	61 y 55	Asp	tac Tyr	19
		60)	120	A3 E	, rne	65	Tyr	Ile	Gln	Ser	Leu 70	Asn	Thr	Pro	243
- 3 -	75	O1 u	GIII	ALG	nis	80	Asn	Arg	Gln	Ala	61 u 85	Leu	Ile	Met	caa Gln	29
90	9		Deu	Leu	95		гàг	Met	Glu	Ala 100	Ile	Gln	Gln	Leu	Glu 105	339
				110	OII.	tac Tyr	Leu	GIY	115	ser	Tyr	Phe	Phe	Gln 120	Asp	387
		-,-	125	116	Leu	agt Ser	ser	130	His	Asn	Glu	Pro	Arg 135	Tyr	Phe	435
		140	nap	Leu	TYE	ttc Phe	145	Ala	Leu	Gly	Phe	Arg 150	Ile	Leu	Arg	483
	155					tcc Ser 160	GIU	мар	vaı	rne	165	Cys	Phe	Lys	Ile	531
70	-			աբ	175	aat Asn	ALA .	ASN	Leu .	180	GIn	Asp	Thr	Lys	Gly 185	579
tg	tta	caa	ctt	tat	gaa	gca	tct	ttc	ctt	ttg	aga	gaa	ggt	gaa	gat	627

Met	Leu	Gln	Leu	Tyr 190		Ala	Ser	Phe	Leu 195	Leu	Arg	Glu	Gly	Glu 200	Asp	
aca Thr	ttg Leu	gag Glu	Cta Leu 205	gca Ala	aga Arg	cga Arg	ttt Phe	tcc Ser 210	acc Thr	aga Arg	tct Ser	cta Leu	cga Arg 215	gaa Glu	aaa Lys	675
ttt Phe	gat Asp	gaa Glu 220	Gly	ggt	gat Asp	gaa Glu	att Ile 225	gat Asp	gaa Glu	gat Asp	cta Leu	tca Ser 230	tcg Se <i>r</i>	tgg Trp	att Ile	723
cgc Arg	cat His 235	tcc Ser	ttg Leu	gat Asp	ctt Leu	Pro 240	ctt Leu	cat His	tgg Trp	agg Arg	gtc Val 245	caa Gln	gga Gly	tta Leu	gag Glu	771
gca Ala 250	aga Arg	tgg Trp	ttc Phe	tta Leu	gat Asp 255	gct Ala	tat Tyr	gcg	agg Arg	agg Arg 260	ccg Pro	gac Asp	atg Met	aat Asn	cca Pro 265	819
Leu	Ile	Phe	Lys	Leu 270	Ala	Lys	Leu	Asņ	Phe 275	aat Asn	Ile	Val	Gln	Ala 280	Thr	867
tat Tyr	Caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 290	agg Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	tcg Ser	tgc Cys	915
Leu	Ala	Glu 300	Lys	Leu	Pro	Phe	Val 305	Arg	Asp	agg Arg	Ile	Val 310	Glu	Cys	Phe	963
ttt Phe	tgg Trp 315	gcc Ala	atc Ile	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	aga Arg	1011
330 Lys	Met	Ala	Ala	Val	11e 335	Ile	Thr	Phe	Ile	aca Thr 340	Ile	Ile	Asp	Asp	Val 345	1059
Tyr	qzA	Val	Tyr	Gly 350	Thr	Ile	Glu	Glu	Leu 355	gaa Glu	Leu	Leu	Thr	Asp 360	Met	1107
att Ile	cgc Arg	aga Arg	tgg Trp 365	gat Asp	aat Asn	aaa Lys	tca Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	cca Pro	tat Tyr 375	tat Tyr	atg Met	1155
caa Gln	gtg Val	tgc Cys 380	tat Tyr	ttg Leu	gca Ala	cta Leu	tac Tyr 385	aac Asn	ttc Phe	gtt Val	tct Ser	gag Glu 3 90	Arg	gct Ala	tac Tyr	1203
gat Asp	att Ile 395	cta Leu	aaa Lys	gat Asp	caa Gln	cat His 400	ttc Phe	aac Asn	agc Ser	atc Ile	cca Pro 405	tat Tyr	tta Leu	cag Gln	a ga Arg	1251
tcg Ser 410	tgg Trp	gta Val	agt Ser	ttg Leu	gtt Val 415	gaa Glu	gga Gly	tat Tyr	ctt Leu	aag Lys 42 0	gag Glu	gca Ala	tac Tyr	tgg Trp	tac Tyr 425	1299
tac	aat	ggc	tat	aaa	cca	agc	ttg	gaa	gaa	tat	ctc	aac	aac	gcc	aag	1347

PCT/US98/20120 WO 99/15624

Tyr	Asn	Gly	Tyr	Lys 430	Pro	Ser	Leu	Glu	Glu 435	Tyr	Leu	Asn	Asn	Ala 440	Lys	
att Ile	tca Ser	ata Ile	tcg Ser 445	gct Ala	cct Pro	aca Thr	atc Ile	ata Ile 450	tcc Ser	cag Gln	ctt Leu	tat Tyr	ttt Phe 455	aca Thr	tta Leu	1395
gca Ala	aac Asn	tcg Ser 460	att Ile	gat Asp	gaa Glu	aca Thr	gct Ala 465	atc Ile	gag Glu	agc Ser	ttg Leu	tac Tyr 470	caa Gln	tat Tyr	cat His	1443
aac Asn	ata Ile 475	ctt Leu	tac Tyr	cta Leu	tca Ser	gga Gly 480	acc Thr	ata Ile	tta Leu	agg Arg	ctt Leu 485	gct Ala	gac Asp	gat Asp	ctt Leu	1491
ggg Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1539
cag Gln	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	gag Glu	aga Arg	gag Glu	gcg Ala	gtg Val 520	gaa Glu	1587
cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1635
aca Thr	aca Thr	gcc Ala 540	agc Ser	gat Asp	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat Asp	ttg Leu	gtt Val 550	gcg Ala	gcc Ala	gca Ala	1683
gct Ala	aat Asn 555	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	Gly ggg	gat Asp	gjå ååå	1731
cat His 570	ggc Gly	gtg Val	caa Gln	cac His	tca Ser 575	gaa Glu	ata Ile	cat His	caa Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	ctg Leu	cta Leu 585	1779
ttc Phe	cag Gln	cct Pro	tat Tyr	gtct	gaat	aa a	tcga	aaat	c ca	acct	acta	tgt	atco	ctc		1831
gata	atat	at t	cttg	gggt	t aa	catg	rttta	att	aaag	ittc	taat	taaa	ag a	ıgctg	aatcg	1891
atco	tcaa	aa a	aaaa	aaaa	a											1911
<210	> 6															

<211> 589 <212> PRT <213> Salvia officinalis

<400> 6

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro

35 His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80 Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190 Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys

355 360 365 Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430 Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Sor Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Scr Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr

<210> 7

<211> 2022

<212> DNA

<213> Salvia officinalis

<400> 7

gcaattcaag ccaaaaatca aagggaaaaa aattgtacct tcttaaaatg tgtagctttg 60
gaatgcaaat ggcgtttcct agcaacccaa ttaaacatct tcataactca gacatcaagt 120
cttcaaaact aatttccagt aatagaattg gtagcagtga tgctgctcgc ctccgcctgc 180
attgctcctc gcagcaacac ggtgccgatg agctccaaac gcaccgacga tctggaaact 240
acagtccttc ccgatggat ttcgattata ttcaatcact ccacagtgat tataaggaag 300

anagacacag aagaagggct agtgagctag ttatgcaagt gaagaagcta atagagaaag 360 $_{f}$. aaacggatcc cactcgacag ttggagctaa tggatgactt gcagaggctg ggcctgggtg 420 atcatttcca ggatgaattc aaggasatct taatctctgt atatttggac sataaatatt 480 acaagagtaa tgtggataat atgaaaaagg ccgaaaggga tttgtactcg acggctcttg 540 cattcagact cottagacaa catggtttto atgttgctcc agaggtgttg gggtgtttca 600 agaacgatga gggcgacttc gaaccaagcc ttgtccatga caccagagga ttgctgcaac 660 tgtacgaagc ttccttcttg ctgacacaag gcgaaaacac actcgagtta gctagagaat 720 ttgcctccag aattctgcag gagaaactac tgaatgatga gattgatgac attaaccttt 780 cgacatggat actcaattct ttggacatcc caatccattg gaggattgaa agggtgaaca 840 caagtgtgtg gatagaagca tacaagaggc gagccgacat gaatccaaca gtgctggatc 900 ttgccatact ggacaccaat attgtacaag cacagtatca ggaggaactc aaacagaact 960 tacagtggtg gagaaattca ggaattgtgg agaagettee ettegtgagg aacaggetag 1020 tggagtccta cttttggagc gttgggatcg tgcagcctcg tcaacatgga attggaagaa 1080 tggcattggg caaatccatc gctcttataa caaccataaa tgatgtttat gatgtgtatg 1140 gtacattaga agaactcgaa caattcacag acgtcattcg aagatgggat ataagttcaa 1200 tagacaaact coctagetat atgcaactgt gttttettge actgcacaac tttgtgaacg 1260 atacggccta tgatgtgcta aaagagcaag gtttcaacat catcccatat ctccgaaaat 1320 cgtggatgga tttggtggag gcatatctgg tggaggccaa gtggtaccac agtggataca 1380 aaccaaatet ggaagagtat ttggagaaet catggatete agaeteagge eetgetgtae 1440 tagcccaagc atttttcggc gtaacacatt ctcttacaga ggaggccgtc cacagtttgt 1500 acggacacca cgatttaatt cgttcgtcat caatgatttt gcgacttgct gatgatctag 1560 gaacetette ggaatgggee atgtgaaaeg ggacagteea atttggaaag tgggeeatgt 1620 gaaacgagac ggagggagta atacatcaac aaatcaacac ttgcttcttc caccctgcaa 1680 cactctaget acgtacetat gtatatatta tatatgcata tgcattgctt gcacacatta 1740 atcaaggaat aatcaatgca tcaccatata tatctacttc tattttatat gttctacttc 1800 agctgttttt actgtcttta attcactata aacaaatatt gcgtatattt tcgagaatgg 1920 aattaataac atgatttttg agaaaaaaa tgaaattatg taggaattaa agataaaatt 1980 tgaaaaaaaa aaaaaaaaa ctcgagggg gcccgtaccc aa 2022

```
<210> 8
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: PCR primer 1F
 <220>
 <221> misc_feature
 <222> (1)..(26)
 <223> PCR primer 1F where "N' is inosine
 <400> 8
 aaraaygarr arggnganta yaarga
                                                                     26
 <210> 9
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: primer 2F
 <220>
 <221> misc_feature
<222> (1)..(17)
<223> PCR primer 2F where "N" is inosine
<400> 9
ytncarytnt aygargo
                                                                    17
<210> 10
<211> 24
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: primer 3R
<220>
<221> misc_feature
<222> (1)..(24)
<223> PCR primer 3R where "N" is inosine
<400> 10
ctrgtyrang gnmtratrta cgty
                                                                   24
<210> 11
<211> 659
<212> DNA
<213> Salvia officinalis
<400> 11
ttgcatgcct gcaggtcgac tctagaggat ctactagtca tatggattgg tctagattgc 60
atgtaatcgg ggaggttgtc gattgaactc atatcccatc ttcgaatgac gtctgtgaat 120
```

```
tgttccagtt cttccaaagt accatacaca tcgtaaacat catctatcgt ggttataaga 180 p
 gcattcactt tgcccaccgt tatccttgca tttgcatgct gacgaggctg cacgatccca 240
 gtagtccaga agtagcattc cacaagtcgg tccctcgcga aggggagctt ctccacgatg 300
 cacqtacttc tecaccactg taagteetgt ttgagtteet ettgatattg tgettgtaca 360
 atattcgagt ccagtatggc aagctccaac acgatggatt catgtcggat ctcctcttat 420
 aggeategat ccaegtgett geattteete ttecaateet ccaatggatt gggatgteca 480
 aagcatagcg tatcgacaat aaaaggtaaa ggtcttcgtt actttcctcc ttcagtttgt 540
 cctcaaggat tttggcggcg aactccctgc ctaaatccag cgtcttctca ccttccatcg 600
 tcaaaaacga agettegtae aattgeagea taateggate eeeggtaeeg agetegaan 659
 <210> 12
 <211> 584
 <212> DNA
<213> Salvia officinalis
<400> 12
ggtctagact tcatgtagtc ggggaggttg tctattgaac ttatatccca tcttcgaatg 60
acgtotgtga actgttcgag ttottotaat gtaccataca catcataaac atcatttatg 120
gttgttataa gagcgatgga tttgcccaat çccattcttc caattccatg ttgacgaggc 180
tgcacgatcc caacgeteca amagtaggae tecaetagee tgtteeteme gaagggaage 240
ttctccacaa ttcctgaatt tctccaccac tgtaagttct gtttgagttc ctcctgatac 300
tgtgcttgta caatattggt gtccagtatg gcaagatcca gcactgttgg attcatgtcg 360
gctcgcctct tgtatgcttc tatccacaca cttgtgttca ccctttcaat cctccaatgg 420
attgggatgt ccaaagaatt gagtatccat gtcgaaaggt taatgtcatc aatctcatca 480
ttcagtagtt teteetgeag aattetggag gcaaattete tagetaaete gagtgtgttt 540
tegeettgtg teageaagaa ggaageeteg tacaactgca geat
                                                                  584
<210> 13
<211> 6
<212> PRT
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: conserved
     amino acid sequence
<220>
<221> DOMAIN
<222> (1)..(6)
<223> Conserved sequence found in all
```

```
prenyltransferases. Xaa at position 1 is I, L or
       V. Xaa at positions 4 and 5 represent any amino
       acid
 <400> 13
 Xaa Asp Asp Xaa Xaa Asp
 <210> 14
 <211> 5
 <212> PRT
 <213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: conserved
       sequence
<220>
<221> DOMAIN
 <222> (1)..(5)
<223> Conserved prenyltransferase sequence. Xaa at
      position 4 represents any amino acid. Xaa at position 5 represents Glu or Asp
<400> 14
Cys Tyr Met Xaa Xaa
<210> 15
<211> 10
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: active site
      peptide
<220>
<221> DOMAIN
<222> (1)..(10)
<223> Monoterpene synthase active site peptide
<400> 15
Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu
 1
<210> 16
<211> 2024
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:
      computer-generated nucleic acid encoding
      (+)-bornyl diphosphate synthase
<220>
```

PCT/US98/20120 WO 99/15624

<221> CDS <222> (11)..(1804)

<400> 16 gatcacaaaa atg tct atc att agc atg aac gta tcg atc ctt agc aag Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys 49 cca cta aat tgc ctc cac aac ttg gag agg aga cct tca aaa gcc ttg Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu 20 ett gëc eet tge aet gea eee aee get ege ete egg gea tet tge tee Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser tca aaa cta caa gaa gct cat caa atc cga cga tct gga aac tac caa 193 Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln cct gcc att tgg gat tcc aat tac att cag tct ctc aat act cca tat Pro Ala Ile Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr acg gag gag agg cac ttg gat aga aaa gca gag ctg att gtg caa gtg 289 Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val agg ata ctg cta aag gaa aaa atg gag cct gtt caa caa ttg gag ttg 337 Arg Ile Leu Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu 385

att cat gac ttg aaa tat ttg ggg ctc tcg gat ttt ttt caa gat gag Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu

att aag gag atc tta ggt gtt ata tac aat gag cac aaa tgc ttt cac 433 Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His 135

aat aat gaa gta gag aaa atg gat ttg tat ttc aca gct ctt gga ttc 481 Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe 150

aga etc etc aga caa cat ggt ttt aat att tec caa gat gta ttt aat 529 Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn 165

tgt ttc aag aac gag aag ggt att gat ttc aag gca agc ctt gct caa 577 Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln 180

gat acg aag gga atg tta caa ctg tat gaa gcg tct ttc ctt ttg aga 625 Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg

aaa ggt gaa gat aca ttg gag ctt gca aga gaa ttt gcc aca aaa tgt 673 Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys

L	tg (cag Sln	aa Ly	a aa 3 Ly 22	a ci s Le S	tt g ≘u A	at d sp (gaa Slu	G1,	gg Gl 23	yм	at sn	gaa Glu	a at	tt d Le i	gat Asp	ga G1 23	u A	at sn	cta Leu	721
Le	a 1 eu 1	tg Leu	Tri 240	g at p Il p	t co e Aı	gc ca	ac t is S	ct	ttg Leu 245	AS,	t c p L	tt eu	ect Pro	ct Le	u F	ac is 250	tg Tr	g a	gg gg	att Ile	769
G1	n s	er 55	gta Va]	ga Gl	g go	a aç a Aı	.y .	gg E p	ttc Phe	at:	aga A	at sp	gct Ala	ta Ty 26	'I A	rcg Lla	ag. Ar	a ag g Ai	:g	cca Pro	817
ga As 27	ca pM 0	tg et	aat Asn	Pre	a ct o Le	t at u Il 27	. C	tc he	gag Glu	ct: Le:	g g g	La	aaa Lys 280	Le	ca uA	ac sn	tte Phe	c as a As	t	att Ile 285	865
at Il	t c e G	aa ln	gca Ala	aca Thi	a ca Hi 29		a ca n G	aa ln	gaa Glu	Cto	aa Ly 29	73 .	gat Asp	ct. Lei	c t u S	cg er	ago Aro	g tg g Tr 30	p	tgg Trp	913
agt Se:	t a	ga rg	tta Leu	tgc Cys 305	tte Phe	c cc	t ga o Gl	sa a iu 1	aag Lys	Ctc Leu 310	PI	a i	ttt Phe	gto Val	g ad L A	gg	gat Asp 315	Ar	g g	ctc Leu	961
gtt Va]	E ga	ia i	tcc Ser 320	ttc Phe	ttt Phe	tg:	g gc		tt /al 825	GJ A āāā	at Me	g t	tt ?he	gaç Glu	g co 1 Pı 33		cat His	Ca. Gl:	a d	cat His	1009
gga	ta Ty 33	r (sag Sln	aga Arg	aaa Lys	ato Met	g gc : Al : 34	a 7	Ja Icc	aca Thr	at Il	t a e I	le	gtt Val 345	. Le	a e	gca Ala	aca Th	a (gtt /al	1057
350		-	•		tac Tyr	355			ĀΤ	GIA	Thi	3	60	Asp	G1	u l	Leu	Glı	1 I	eu 165	1105
			•		ttt Phe 370	-,0	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	y I	ı,	мзр	375	5	Iu	Ser	11	e 7	hr	Arg 380	L	eu	1153
	_		-	385	caa Gln		Cy.	• • :	Y	390	сту	ν,	al I	Hls	As	n I	yr 95	Ile	S	er	1201
		4	00	•	gat Asp		200	4(5	31 U	nıs	G.	Ly I	Phe	410	e C	ys	Leu	G	ln	1249
	415	•		•	tcg Ser		420)	.p .	.eu	val	G	4	125	туі	: P.	he	His	G.	Lu	1297
430			•		cac His	435	GLY	1 y	. I	nr :	PIO	44	0	eu	Asp	G.	lu '	Tyr	Le	eu 15	1345
aac Asn	atc Ile	gc Al	c a a L	-	att Ile 450	tca Ser	gtg Val	gc	g t a S	er i	ect Pro 155	gc A 1	a a a I	ta le	ata Ile	to Se	er	cca Pro 460	a c Th	ic ir	1393

PCT/US98/20120 WO 99/15624

Tyr	Phe	aca Thr	Phe 465	gca Ala	aac Asn	gcg Ala	tct Ser	cat His 470	gac Asp	aca Thr	gca Ala	gtc Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	caa Gln 480	TYE	cat His	gac Asp	ata Ile	ctt Leu 485	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile 490	att Ile	ttg Leu	agg Arg	1489
ctt Leu	Pro 495	gac Asp	gat Asp	ctt Leu	el A aaa	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	gcg Ala	aga Arg	ggc Gly	gac Asp	1537
gtg Val 510	ecg Pro	aaa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	aca Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585
gag Glu	gag Glu	gcg Ala	gtg Val	gag Glu 530	cac His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540	aag Lys	1633
gat Asp	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	Ala	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	ggt Gly	atg Met	1681
gtg Val	gcg Ala	ggc Gly 560	gca Ala	gct Ala	aat Asn	TIE	ggg Gly 565	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	att Ile	tat Tyr	ctc Leu	1729
	gga Gly 575	gat Asp	ej À aaa	ttt Phe	Gly ggc	gtg Val 580	caa Gln :	cac His	tcg Ser	Lys	acg Thr 585	tac Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 590	Gl y ggc	cta Leu	ctg Leu	rne	gag Glu 1 595	cct Pro '	tat (Tyr)	gca Ala	tgaa	caaa	tg g	gaga	ctgc	t		1824
tgat	atat	at t	aatti	tggc.	a ca	caa	taat	tgc	atgti	tat a	atat	gttg	ga a	aataa	agtgt	1884
etgg	ttga	ga t	gtcai	tgtg	g tg	tatta	atct	aaa	taati	tca a	aggt	tgcct	t g	tttai	gtag	1944
cgg	tggt	gc aa	actad	ctc	c cat	tcaa	aatc	aati	caaat	tct a	aaac	agtc	ga gt	caaç	gctcg	2004
get	=gag	ga aa	aaaa	aaaa	a											2024
210:	598															

<213> Artificial Sequence

<400> 17

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40 45

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Ile 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys
165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Leu Trp
225 230 235 240

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 245 250 255

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 265 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 290 295 300

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 305 310 315 320

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 340 345 350

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355 360 365

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 370 380

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp
420 425 430

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala
435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 18

<211> 2024

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid encoding (+)-bornyl diphosphate synthase

<220>

<221> CDS

<222> (11)..(1804)

	00>															
ga	itca (:aaaa	atg Met 1	tct Ser	atc Ile	att Ile	agc Ser 5	atg Met	aac Asn	gta Val	tcg Ser	atc Ile 10	ctt Leu	ago Ser	aag Lys	49
Pr		a aa u As .5	t tg: n Cy:	c cto s Leo	c cac u His	aad Asr 20	rei	g gaq ı Glı	g ago	g ag	a cc g Pro 2	o Se	a aa r Ly:	a go	c ttg a Leu	97
ct Le 3		c cc l Pr	t tgo o Cys	c act	gca Ala 35	PEC	aco Thi	gct Ala	cgo Arc	Cto Let 40	ı Arç	g gca	a tci a Sei	t tg r Cy	c tcc s Ser 45	145
	- - ,	. .	u 911.	50)	nıs	GIT	lle	Arg 55	Arg	, Ser	: Gl	/ Asr	1 Ty	-	193
			65	, wan) Set	Asn	туг	70	GIN	Ser	Leu	As.	75	Pr	a tat o Tyr	241
		80)	urs	Leu	Asp	85	гÀ2	Ala	Glu	Leu	90	Val	Gli	a gtg n Val	289
3	95	5	. Deu	пуз	GIU	100	Met	GIU	Pro	Val	Gln 105	Gln	Leu	Glu	ttg Leu	337
110			Leu	Lys	tat Tyr 115	ren	GIĀ	Leu	Ser	Asp 120	Phe	Phe	Gln	Asp	Glu 125	385
			-10	130	ggt Gly	Val	TTE	Tyr	135	Glu	His	Lys	Cys	Phe 140	His	433
		024	145	GIU	aaa Lys	Met	Asp	150	Tyr	Phe	Thr	Ala	Leu 155	Gly	Phe	481
•		160	•••9	GIN	cat His	GIĀ	165	ASN	ITE	Ser	Gln	Asp 170	Val	Phe	Asn	529
	175	•				180	116	мэр	rne	гÀз	185	ser	Leu	Ala	Gln	577
190		•	1		tta (Leu (195	3411 <i>1</i>	Dea	IÿE (siu .	200	Ser	Phe	Leu	Leu	Arg 205	625
			•	210	ttg (Leu (·	ueu /	-ta /	215	GIU	rne .	ATa	Thr	Lys 220	Cys	673
Leu	cag Gln	-	aaa Lys 225	ctt (Leu)	gat o	gaa g Slu (TA (ggt a Gly A 230	aat (Asn (gaa : Slu :	att d	Asp	gag Glu 235	aat Asn	cta Leu	721

tt Le	a tt	g to su Ti 24		t co le Ai	gc ca Fg Hi	ic to is Se	t tt r Le 24	u As	t ct p Le	et co	ct ct co Le	c ca u Hi 25	s Tr	g ag	gg at	t 769 _, .
ca Gl	a ag n Se 25	r Va 5	a ga 1 G1	ig go .u Al	a ag la Ar	a tg g Tr 26	Б ħu	c at e Il	a ga e As	it go ip Al	t ta a Ty 26	r Al	g aç	a ag g Ar	g cca	a 817
ga As 27		g aa t As	t co n Pr	a ct o Le	t at u Il 27	e rn	c ga e Gl	g ct u Le	t gc u Al	c aa a Ly 28	's Le	c aa u Ası	tto n Pho	c aa e As	t att n Ile 285	•
at: Il	t ca e Gl	a gc n Al	a ac a Th	a ca r Hi 29	2 61	a ca n Gl:	a gaa n Glu	a cto	g aa u Ly 29	s As	t ct p Le	c tco	age Are	tg Tr 30	g tgg P Trp O	913
agt Se	t ag	a tt g Le	a tg u Cy: 30:	o FM	c cc e Pr	t gaa	a aaq u Lys	g cto Let 310	ı Pro	a tt o Ph	t gte e Val	g agg	gat Asp 315	Ar	g ctc g Leu	961
gtt Val	gaa LG1	Se:		tti Pho	t tgg	g gcç	g gtt Val 325	. Gry	g ato	g tti E Phe	t gag e Glu	g cca 1 Pro 330	His	Caa Glr	a cat 1 His	1009
gga Gly	Ty:		g aga n Arg	a aaa g Lys	a ato Met	g gcc Ala 340	NT a	aca Thr	att	att	t gtt Val 345	. Leu	gca Ala	aca Thr	gtt Val	1057
350				: 1 <u>y</u> 1	355	, vai	Tyr	GIĀ	Thr	360	ı Asp	Glu	Leu	Glu	cta Leu 365	1105
		1.00		370	, TAS	Arg	Trp	Asp	375	Glu	Ser	Ile	Thr	Arg 380		1153
	-3-	-32	385	GIN	Leu	Cys	Tyr	390	GIY	Val	His	Asn	Tyr 395	Ile		1201
•		400	- 4-	Asp	116	Leu	105	GIU	Hls	Gly	ttc Phe	Phe 410	Cys	Leu	Gln	1249
_	415		-,-		,,,,	420	vsb	ren	vai	GIU	gca Ala 425	Tyr	Phe	His	Glu	1297
430	_	-	-4-		435	Cly	TYL	Int	PEO	440	ctg Leu	Asp	GLu	Tyr	Leu 445	1345
				450		141	AL a	Ser	455	ALA	ata Ile	Ile	Ser	Pro 460	Thr	1393
cat Cyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	Jer	cat His 470	gac Asp	aca Thr	gca Ala	Val	atc (Ile / 475	gac Asp	agc Ser	1441

ttg Leu	tac Tyr	caa Gln 480	tat Tyr	cat His	gac Asp	ata Ile	ctt Leu 485	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile 490	att Ile	ttg Leu	agg Arg	1489
ctt Leu	CCC Pro 495	gac Asp	gat Asp	ctt Leu	ej A ada	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	gcg Ala	aga Arg	ggc	gac Asp	1537
gtg Val 510	ccg Pro	aaa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	aca Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585
gag Glu	gag Glu	gcg Ala	gtg Val	gag Glu 530	cac His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540	aag Lys	1633
gat Asp	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	gcc Ala	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	ggt Gly	atg Met	1681
gtg Val	gcg Ala	ggc Gly 560	gca Ala	gct Ala	aat Asn	att Ile	ggg G1 y 565	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	att Ile	tat Tyr	ctc Leu	1729
****	gga Gly 575	gat Asp	ej A aaa	ttt Phe	Giy	gtg Val 580	caa Gln	cac His	tcg Ser	Lys	acg Thr 585	tac Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 590	GJ Y ggc	cta Leu	ctg Leu	Phe	gag Glu 595	cct Pro	tat Tyr	gca Ala	tgaa	caaa	tg g	gaga	ctgc	:t		1824
tgat	atat	at t	aatt	tggc	a ca	ccaa	taat	tgc	atgt	tat	atat	gttg	ga a	aata	agtgt	1884
ctgg	ttga	ga t	gtca	tgtg	g tg	tatt	atct	aaa	taat	tca	aggt	tgcc	tt g	ttta	tgtag	1944
ccgg	tggt	gc a	acta	cctc	c ca	ttca	aatc	aat	taaa	tct .	aaac	agtc	ga g	tcaa	gctcg	2004
agct	cgag	ga a	aaaa	aaaa	a											2024

<210> 19 <211> 598 <212> PRT

<213> Artificial Sequence

<400> 19

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Leu Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205 Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 250 Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 360 Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg 405 410 Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly **Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp** Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu Leu Phe Glu Pro Tyr Ala 595 <210> 20 <211> 2024 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence encoding (+)-bornyl diphosphate synthase <220> <221> CDS <222> (11) .. (1804) <400> 20 gatcacaaaa atg tct atc att agc atg aac gta tcg atc ctt agc aag 49 Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys

			1				5					10				
cca Pro	cta Leu 15	aat Asn	tgc Cys	ctc Leu	cac His	aac Asn 20	ttg Leu	gag Glu	agg Arg	aga Arg	cct Pro 25	tca Ser	aaa Lys	gcc Ala	ttg Leu	97
ctt Leu 30	gtc Val	cct Pro	tgc Cys	act Thr	gca Ala 35	ccc Pro	acc Thr	gct Ala	cgc Arg	ctc Leu 40	cgg Arg	gca Ala	tct Ser	tgc Cys	tcc Ser 45	145
tca Ser	aaa Lys	cta Leu	caa Gln	gaa Glu 50	gct Ala	cat His	caa Gln	atc Ile	cga Arg 55	cga Arg	tct Ser	gga Gly	aac Asn	tac Tyr 60	Ca a Gl n	193
cct Pro	gcc Ala	ctt Leu	tgg Trp 65	gat Asp	tcc Ser	aat Asn	tac Tyr	att Ile 70	cag Gln	tct Ser	atc Ile	aat Asn	act Thr 75	cca Pro	tat Tyr	241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	ttg Leu	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	gtg Val	caa Gln	gtg Val	289
agg Arg	ata Ile 95	ctg Leu	cta Leu	aag Lys	gaa Glu	aaa Lys 100	atg Met	gag Glu	cct Pro	gtt Val	caa Gln 105	caa Gln	ttg Leu	gag Glu	ttg Leu	337
Ile 110	His	Asp	Leu	Lys	Tyr 115	Leu	Gly	Leu	tcg Ser	120	Phe	rne	GIN	Азр	125	385
Ile	Lys	Glu	Ile	Leu 130	Gly	Val	Ile	Tyr	aat Asn 135	Glu	Hls	Lys	Cys	140	HIS	433
Asn	Asn	Glu	Val 145	Glu	ГÀЗ	Met	Asp	Leu 150	tat Tyr	Phe	Thr	ALA	155	GTÀ	Pne	481
Arg	Leu	Leu 160	Arg	Gln	His	Gly	Phe 165	Asn	att Ile	Ser	Gln	170	Val	Pne	ASn	529
tgt Cys	ttc Phe 175	aag Lys	aac Asn	gag Glu	aag Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	Ser	ctt Leu	gct Ala	Caa Gln	577
gat Asp 190	Thr	aag Lys	gga Gly	atg Met	tta Leu 195	caa Gln	ctg Leu	tat Tyr	gaa Glu	gcg Ala 200	Ser	ttc Phe	ctt Leu	ttg Leu	aga Arg 205	625
aaa Lys	ggt G1y	gaa Glu	gat Asp	aca Thr 210	Leu	gag Glu	ctt Leu	gca Ala	aga Arg 215	Glu	ttt Phe	gcc	aca Thr	aaa Lys 220	Cys	673
ctg Leu	cag Gln	aaa Lys	aaa Lys 225	Leu	gat Asp	gaa Glu	ggt Gly	ggt Gly 230		gaa Glu	att	gat Asp	gag Glu 235	Asn	cta Leu	721
tta Leu	ttg Leu	tgg	att	cgc Arg	cac His	tct Ser	ttg Leu	gat Asp	ctt Leu	Pro	ctc Leu	cac His	tgg	agg Arg	att	769

240	245	250	
caa agt gta gag gca	aga too tto ata gat	get tat geg aga agg	-c-a

caa Glr	agt Ser 255	. A91	gag Glu	gca Ala	aga Arg	tgg Trp 260	Phe	ata Ile	gat Asp	gct Ala	tat Tyr 265	gcg Ala	aga Arg	Arg	Pro	817
gac Asp 270	Met	aat Asn	cca Pro	Leu	Ile 275	ttc Phe	gag Glu	ctt Leu	gcc	Lys 280	Leu	aac Asn	ttc Phe	aat Asn	att Ile 285	865
att Ile	caa Gln	gca Ala	aca Thr	cat His 290	Caa Gln	caa Gln	gaa Glu	ctg Leu	aaa Lys 29 5	Asp	ctc Leu	tcg Ser	agg Arg	tgg Trp 300	tgg Trp	913
agt Ser	aga Arg	tta Leu	tgc Cys 305	ttc Phe	Pro	gaa Glu	aag Lys	ctc Leu 310	cca Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	agg Arg	ctc Leu	961
gtt Val	gaa Glu	tcc Ser 320	ttc Phe	ttt Phe	tgg Trp	gcg Ala	gtt Val 325	ej Å aaa	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
gga	tat Tyr 335	cag Gln	aga Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Thr	att Ile	att Ile	gtt Val 345	tta Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	gat Asp	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt Gly	aca Thr	cta Leu 360	gat Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac Asp	acg Thr	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	ctt Leu	1153
ect Pro	tat Tyr	tac Tyr	atg Met 385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 390	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
gat Asp	gca Ala	gca Ala 400	tat Tyr	gat Asp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caa Gln	1249
tat Tyr	ctc Leu 415	cgg Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat As p	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297
gca Ala 430	aag Lys	tgg Trp	tac Tyr	cac His	agc Ser 435	ggt Gly	tat Tyr	aca Thr	cca Pro	agc Ser 440	ctg Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
aac Asn	atc Ile	gcc Ala	aag Lys	att Ile 450	tca Ser	gtg Val	gcg Ala	tct Ser	cct Pro 455	gca Ala	ata Ile	ata Ile	tcc Ser	cca Pro 460	acc Thr	1393
tat Tyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	tct Ser	cat His 470	gac Asp	aca Thr	gca Ala	Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	caa Gln	tat Tyr	cat His	gac Asp	ata Ile	ctt Leu	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile	att Ile	ttg Leu	agg Arg	1489

		480					485					490				
ctt Leu	ccc Pro 495	Asp	gat A sp	ctt Leu	el A aaa	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	gcg Ala	aga Arg	G1 y ggc	gac Asp	1537 ∱
gtg Val 510	ccg Pro	aaa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	aca Thr	aat Asn	gct Ala	agt Ser	geg Glu 525	1585
gag Glu	gag Glu	gcg Ala	gtg Val	gag Glu 530	cac His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540	aag Lys	1633
gat Asp	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	gcc Ala	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	ggt Gly	atg Met	1681
gtg Val	gcg Ala	ggc Gly 560	gca Ala	gct Ala	aat Asn	att	ggg Gly 565	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	att Ile	tat Tyr	ctc Leu	1729
urs	gga Gly 575	gat Asp	Gly ggg	ttt Phe	ej ggc	gtg Val 580	caa Gln	cac His	tcg Ser	aaa Lys	acg Thr 585	tac Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 59 0	ggc Gly	cta Leu	ctg Leu	ttc Phe	gag Glu 595	cct Pro	tat Tyr	gca Ala	tgaa	caaa	itg g	gaga	ctgo	:t		1824
tgat	atat	at t	aatt	tggc	a ca	ccaa	taat	tgc	atgt	tat	atat	gttg	ga a	aata	agtgt	1884
ctgg	ttga	iga t	gtca	tgtç	g to	tatt	atct	: aaa	taat	tca	aggt	tgcc	tt g	ttta	tgtag	1944
ccgg	tggt	gc a	acta	ccto	c ca	ttca	aato	aat	taaa	tct	aaac	agto	ga g	tcaa	gctcg	2004
agct	cgag	ga a	aaaa	aaaa	ıa											2024
<210: <211: <212: <213:	> 59 > PR	8 T	cial	Seq	(uenc	:e										
<400																
let :	Ser	Ile	Ile	Ser 5	Met	Asn	Val	Ser	Ile 10	Leu	Ser	Lys	Pro	Leu 15	Asn	
ys I	Leu	His .	Asn 20	Leu	Glu	Arg	Arg	Pro 25	Ser	Lys .	Ala	Leu	Leu 30	Val	Pro	
ys T	Chr .	Ala 35	Pro '	Thr	Ala	Arg	Leu 40	Arg .	Ala	Ser	Суз	Ser 45	Ser	Lys	Leu	
	30					55				Asn '	60					

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 135 Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 305 315 Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 375 Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 390

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 425 Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 455 Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 470 . Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly Ala Ala Ann Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 Leu Phe Glu Pro Tyr Ala 595 <210> 22 <211> 2024 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <222> (11)..(1804) <223> computer-generated nucleic acid sequence encoding (+)-bornyl diphosphate synthase <400> 22 gatcacaaaa atg tot atc att agc atg aac gta tog atc ott agc aag

10

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys

5

cca Pro	cta Leu 15	Asn	tgc Cys	ctc Leu	cac His	aac Asn 20	Leu	gag Glu	agg Arg	aga Arg	Pro 25	Ser	aaa Lys	gcc Ala	ttg Leu	97
ctt Leu 30	AT	Pro	tgc Cys	act	gca Ala 35	Pro	acc Thr	gct Ala	cgc A rg	ctc Leu 40	cgg Arg	gca	tct Ser	tgc Cys	tcc Ser 45	145
tca Ser	aaa Lys	cta Leu	caa Gln	gaa Glu 50	gct Ala	cat His	caa Gln	atc Ile	cga Arg 55	cga Arg	tct Ser	gga Gly	aac Asn	tac Tyr 60	caa Gln	193
cct Pro	gcc Ala	ctt Leu	tgg Trp 65	gat Asp	tcc Ser	aat Asn	tac Tyr	att Ile 70	cag Gln	tct Ser	ctc Leu	aat Asn	act Thr 75	cca Pro	tat Tyr	241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	gtg Val	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	gtg Val	caa Gln	gtg Val	289
agg Arg	ata Ile 95	ctg Leu	cta Leu	aag Lys	gaa Glu	aaa Lys 100	atg Met	gag Glu	cct Pro	gtt Val	caa Gln 105	caa Gln	ttg Leu	gag Glu	ttg Leu	337
att Ile 110	cat His	gac Asp	ttg Leu	aaa Lys	tat Tyr 115	ttg Leu	Gl y ggg	ctc Leu	tcg Ser	gat Asp 120	ttt Phe	ttt Phe	caa Gln	gat Asp	gag Glu 125	385
att Ile	aag Lys	gag Glu	atc Ile	tta Leu 130	ggt Gly	gtt Val	ata Ile	tac Tyr	aat Asn 135	gag Glu	cac His	aaa Lys	tgc Cys	ttt Phe 140	cac His	433
aat Asn	aat Asn	gaa Glu	gta Val 145	gag Glu	aaa Lys	atg Met	gat Asp	ttg Leu 150	tat Tyr	ttc Phe	aca Thr	gct Ala	ctt Leu 155	gga Gly	ttc Phe	481
aga Arg	ctc Leu	ctc Leu 160	aga Arg	caa Gln	cat Kis	ggt Gly	ttt Phe 165	aat Asn	att Ile	tcc Ser	caa Gln	gat Asp 170	gta Val	ttt Phe	aat Asn	529
tgt Cys	ttc Phe 175	aag Lys	aac Asn	gag Glu	aag Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	agc Ser	ctt Leu	gct Ala	caa Gln	577
gat Asp 190	acg Thr	aag Lys	gga Gly	atg Met	tta Leu 195	caa Gln	ctg Leu	tat Tyr	gaa Glu	gcg Ala 200	tct Ser	ttc Phe	ctt Leu	ttg Leu	aga Arg 205	625
aaa Lys	ggt Gly	gaa Glu	gat Asp	aca Thr 210	ttg Leu	gag Glu	ctt Leu	Ala	aga Arg 215	gaa Glu	ttt Phe	gcc Ala	aca Thr	aaa Lys 220	tgt Cys	673
ctg Leu	cag Gln	aaa Lys	aaa Lys 225	ctt Leu	gat Asp	gaa Glu	ggt Gly	ggt Gly 230	aat Asn	gaa Glu	att Ile	gat Asp	gag Glu 235	aat Asn	cta Leu	721
tta Leu	neu	tgg Trp 240	att Ile	cgc Arg	cac His	Ser	ttg Leu 245	gat Asp	ctt Leu	cct Pro	ctc Leu	cac His 250	tgg Trp	agg Arg	att Ile	769

caa Gln	agt Ser 255	gta Val	gag Glu	gca Ala	aga Arg	tgg Trp 260	ttc Phe	ata Ile	gat Asp	gct Ala	tat Tyr 265	gcg Ala	aga Arg	agg Arg	cca Pro	817
gac Asp 270	Met	aat Asn	cca Pro	ctt Leu	att Ile 275	ttc Phe	gag Glu	ctt Leu	gcc Ala	aaa Lys 280	ctc Leu	aac Asn	ttc Phe	aat Asn	att Ile 285	865
att Ile	Caa Gln	gca Ala	aca Thr	cat His 290	caa Gln	caa Gln	gaa Glu	ctg Leu	aaa Lys 295	gat Asp	ctc Leu	tcg Ser	agg Arg	tgg Trp 300	tgg Trp	913
agt Ser	aga Arg	tta Leu	tgc Cys 305	ttc Phe	cct Pro	gaa Glu	aag Lys	ctc Leu 310	cca Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	agg Arg	ctc Leu	961
gtt Val	gaa Glu	tcc Ser 320	ttc Phe	ttt Phe	tgg Trp	gcg Ala	gtt Val 325	Gly ggg	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
gga Gly	tat Tyr 335	cag Gln	aga Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Th <i>r</i>	att Ile	att Ile	gtt Val 345	tta Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	gat Asp	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt Gly	aca Thr	cta Leu 360	gat Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac Asp	acg Th <i>r</i>	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	ctt Leu	1153
cct Pro	tat Tyr	tac Tyr	atg Met 385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 390	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
gat Asp	gca Ala	gca Ala 400	tat Tyr	gat Asp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caa Gln	1249
tat Tyr	ctc Leu 415	egg Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297 [.]
gca Ala 430	aag Lys	tgg Trp	tac Tyr	cac His	agc Ser 435	ggt Gly	tat Tyr	aca Thr	cca Pro	agc Ser 440	ctg Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
aac Asn	atc Ile	gcc Ala	aag Lys	att Ile 450	tca Ser	gtg Val	gcg Ala	tct Ser	cct Pro 455	gca Ala	ata Ile	ata Ile	tcc Ser	cca Pro 460	acc Thr	1393
tat Tyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	tct Ser	cat His 470	gac Asp	aca Thr	gca Ala	gtc Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	caa Gln 480	tat Tyr	cat His	gac Asp	ata Ile	ctt Leu 485	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile 490	att Ile	ttg Leu	agg Arg	1489

cti Le:	Pro 495	Asp	gat Asp	ctt Leu	Gly	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	Ala	aga Arg	ggc	gac Asp	1537
gtç Val 510	ccg Pro	aaa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	Thr	aat Asn	gct Ala	agt Ser	gag Glu 5 25	1585
gaç Glu	gag Glu	gcg Ala	gtg Val	gag Glu 530	cac His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540	aag Lys	1633
gat Asp	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	gcc Ala	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	ggt Gly	atg Met	1681
gtg Val	gcg Ala	ggc Gly 560	gca Ala	gct Ala	aat Asn	att Ile	ggg Gly 565	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	att Ile	tat Tyr	ctc Leu	1729
cac	gga Gly 575	gat Asp	Gly ggg	ttt Phe	ggc Gly	gtg Val 580	caa Gln	Cac His	tcg Ser	aaa Lys	acg Thr 585	tac Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 590	ggc	cta Leu	ctg Leu	ttc Phe	gag Glu 59 5	cct Pro	tat Tyr	gca Ala	tgaa	caaa	ıtg ç	gaga	.ctgd	:t		1824
tga	tatat	at t	aatt	tggc	a ca	ccaa	taat	tgo	atgt	tat	atat	gttg	ga a	aata	agtgt	1884
ctg	gttga	iga t	gtca	tgtg	g tg	tatt	atct	aaa	taat	tca	aggt	tgcc	tt g	rtta	tgta g	1944
ccg	gtggt	gc a	acta	cctc	c ca	ttca	aatc	aat	taaa	tct	aaac	agto	ga g	tcaa	gctcg	2004
agci	tcgag	ga a	aaaa	aaaa	a .											2024
<212 <212	0> 23 l> 59 2> PF 3> Ar	8 T	cial	Seq	uenc	e										
	> 23															
Met 1	Ser	ŢŢĠ	Ile :	Ser 5	Met	Asn '	Val	Ser	Ile 10	Leu	Ser	Lys	Pro	Leu 15	Asn	
Cys	Leu	His .	Asn 1 20	Leu	Glu i	Arg i	Arg	Pro 25	Ser	Lys .	Ala	Leu	Leu 30	Val	Pro	
Cys	Thr	Ala 35	Pro !	Thr .	Alai	Arg 1	Leu 2 40	Arg	Ala	Ser	Cys	Ser 45	Ser	Lys	Leu	
Gln	Glu 50	Ala	His (31n	Ile i	Arg /	Arg :	Ser (Gly .		Tyr 60	Gln	Pro .	Ala	Leu	

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Val Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 120 Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 150 Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 310 Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 345 Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 380 Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp
420
425

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 550 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp
565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585

Leu Phe Glu Pro Tyr Ala 595

<210> 24

<211> 2024

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (11)..(1804)

<223> computer-generated sequence encoding (+)-bornyl diphosphate synthase

<400> 24

gatcacaaaa atg tct atc att agc atg aac gta tcg atc ctt agc aag 49

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys

1 5 10

cca cta aat tgc ctc cac aac ttg gag agg aga cct tca aaa gcc ttg 97 Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu

	1	5				20)				25	5				
cti Lei 30	ı val	e cct L Pro	t tgo Cys	act Thr	gca Ala 35	Pro	acc Thr	gct Ala	cgc Arg	Cto Lev 40	Arg	gca Ala	a tci a Sei	t tgo Cys	tcc Ser 45	:
tca Sei	a aaa Lys	t cta Leu	caa Glr	gaa Glu 50	Ala	cat His	Caa Gln	ato Ile	cga Arg 55	Arg	tct Ser	gga Gl	aac Asr	tac Tyr 60	Gln	19
Pro	gco Ala	: ctt	tgg Trp 65	Asp	tcc Ser	aat Asn	tac Tyr	Ile 70	Gln	s tct Ser	ctc Leu	aat Asn	act Thr 75	Pro	tat Tyr	24
acg Thr	gag Glu	gag Glu 80	Arg	cac His	ttg Leu	gat Asp	aga Arg 85	a aa Lys	gca Ala	gag Glu	gtg Val	att Ile 90	Val	caa Gln	gtg Val	289
AEG	95	ren	Leu	Lys	Glu	aaa Lys 100	Met	Glu	Pro	Val	Gln 105	Gln	Leu	Glu	Leu	337
110	nis	Asp	Leu	Lys	Tyr 115	ttg Leu	Gly	Leu	Ser	Asp 120	Phe	Phe	Gln	Asp	Glu 125	385
116	гàа	GIU	lie	Leu 130	Gly	gtt Val	Ile	Tyr	Asn 135	Glu	His	Lys	Cys	Phe 140	His	433
ASI	Asn	GIU	145	Glu	Lys	atg Met	Asp	Leu 150	Tyr	Phe	Thr	Ala	Leu 155	Gly	Phe	481
Arg	rea	160	Arg	GID	His	ggt Gly	Phe 165	Asn	Ile	Ser	Gln	Asp 170	Val	Phe	Asn	529
cys	175	гÀз	Asn	Glu	Lys	ggt Gly 180	Ile	Asp	Phe	Lys	Ala 185	Ser	Leu	Ala	Gln	577
190	INT	rås	GIÀ	Met	195	caa Gln	Leu	Tyr	Glu	Ala 200	Ser	Phe	Leu	Leu	Arg 205	625
Lys	GIY	GIU	Asp	210	Leu	gag Glu	Leu	Ala	Arg 215	Glu	Phe	Ala	Thr	Lys 220	Cys	673
Deu	GIII	гуя	225	rea	Asp	gaa Glu	GLY	Gly 230	Asn	Glu	Ile	qzA	Glu 235	Asn	Leu	721
204	neu	240	TTE	Arg	Hls	ser	Leu 245	Asp	Leu	Pro	Leu	His 250	Trp	Arg	Ile	· 769
Caa Gln	agt Ser	gta Val	gag Glu	gca Ala	aga Arg	tgg Trp	ttc Phe	ata Ile	gat As p	gct Ala	tat Tyr	gcg Ala	aga Arg	agg Arg	cca Pro	817

	255	5				260	o				265	5				
gac Asp 270	1.100	g aat : Asi	t cca	Let	att 1 11e 275	Phe	gaç Glu	g ctt Lev	gco Ala	280 280	Leu	aac Asn	tto Phe	aat Asn	att Ile 285	865
att Ile	Caa Gln	gca Ala	a aca	Cat His 290	GIN	caa Gln	gaa Glu	ctg Leu	Lys 295	: Asp	cto Leu	tcg Ser	agg Arg	tgg Trp 300		913
agt Ser	aga Arg	tta Leu	tgc Cys 305	Pne	cct Pro	gaa Glu	aag Lys	Leu 310	Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	Arg	ctc Leu	961
gtt Val	gaa Glu	ser 320	Elle	ttt Phe	tgg Trp	gcg	gtt Val 325	ggg Gly	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
GLY	335	GIN	Arg	rys	Met	340	Ala	Thr	Ile	att Ile	Val 345	Leu	Ala	Thr	Val	1057
350	wab	vsb	116	Tyr	355	Val	Tyr	GTA	Thr	cta Leu 360	Asp	Glu	Leu	Glu	Leu 365	1105
	****	vsb	Int	370	Lys	Arg	Trp	Asp	Thr 375		Ser	Ile	Thr	Arg 380	Leu	1153
Pro		TAT	385	GIN	Leu	Cys	Tyr	390	Gly	Val	His	Asn	Tyr 395	Ile	Ser	1201
gat Asp	gca Ala	gca Ala 400	tat Tyr	gat A sp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caz Gln	1249
tat Tyr	ctc Leu 415	cgg Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297
gca a Ala 1 430	aag Lys	tgg T <i>r</i> p	tac Tyr	cac His	agc Ser 435	ggt Gly	tat Tyr	aca Thr	cca Pro	agc Ser 440	ctg Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
aac a Asn 1	atc Ile	gcc Ala	aag Lys	att Ile 450	tca Ser	gtg Val	gcg Ala	tct Ser	cct Pro 455	gca Ala	ata Ile	ata Ile	tcc Ser	cca Pro 460	acc Thr	1393
tat t	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	ser	cat His 470	gac Asp	aca Thr	gca Ala	Val	atc Ile 475	gac Asp	agc Ser	1441
ttg t Leu 1		caa Gln 480	tat Tyr	cat His	gac Asp	тте	ctt Leu 485	tgc Cys	cta Leu	gca Ala	Gly	att Ile 490	att Ile	ttg Leu	ag g ⁻ A rg	1489
ctt c Leu F	ecc (gac Asp	gat Asp	ctt Leu	Gly ggg	aca Thr	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu .	gcg Ala	aga Arg	ggc Gly	gac Asp	1537

	495	5				500)				509	i				
gt (Va) 51(LPEC	j aaa) Lys	aca Thr	atc Ile	cag Gln 515	Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585
gaç Glu	gaç Glu	gcg Ala	gtg Val	gag Glu 530	His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540		1633
gat Asp	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	gcc	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	Gly	atg Met	1681
gtg Val	gcg	ggc Gly 560	gca Ala	gct Ala	aat Asn	att Ile	ggg Gly 565	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	Ile	tat Tyr	ctc Leu	1729
Cac	gga Gly 575	Asp	G] À aàà	ttt Phe	ggc Gly	gtg Val 580	caa Gln	cac His	tcg Ser	aaa Lys	acg Thr 585	Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 590	GIA	cta Leu	ctg Leu	ttc Phe	gag Glu 595	cct Pro	tat Tyr	gca Ala	tgaa	acaa	atg	ggag	actg	ct		1824
tga	tata	tat t	aatt	tgg	ca ca	ecca	ataat	t tgo	atgi	ttat	ata	tgtt	gga a	aaat	aagtgt	1884
ctg	gttg	aga t	gtca	atgt	gg to	gtati	tatci	t aaa	taat	ttca	agg	ttgc	stt (gttt	atgtag	1944
ccg	gtgg	tgc a	acta	acct	cc ca	attca	aato	aat	taaa	atct	aaa	cagto	cga (gtca	agctcg	2004
agc	tcga	gga a	aaaa	aaaa	aa											2024
<21:	0> 2: 1> 5: 2> Pi 3> A:	98 RT	.cial	. Sec	juenc	:e										
	0> 29			_												
nec 1	ser	TTG	IIe	Ser 5	Met	Asn	Val	Ser	Ile 10	Leu	Ser	Lys	Pro	Leu 15	Asn	
Cys	Leu	His	Asn 20	Leu	Glu	Arg	Arg	Pro 25	Ser	Lys	Ala	Leu	Leu 30	Val	Pro	
Cys	Thr	Ala 35	Pro	Thr	Ala	Arg	Leu 40	Arg	Ala	Ser	Суз	Ser 45	Ser	Lys	Leu	
Gln	Glu 50	Ala	His	Gln	Ile	Arg 55	Arg	Ser	Gly	Asn	Туг 60	Gln	Pro	Ala	Leu	
Trp 65	Asp	Ser	Asn	Tyr	Ile 70	Gln	Ser	Leu	Asn	Thr 75	Pro	Tyr	Thr	Glu	Glu 80 -	
Arg	His	Leu	Asp .	Arg 85	Lys	Ala	Glu	Val		Val	Gln	Val	Arg		Leu	
				-					90					95		

100	105	110
-----	-----	-----

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155 160

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 225 230 235 240

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 245 250 255

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 265 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 290 295 300

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 310 315 320

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 340 345 350

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355 360 365

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 370 380

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg 405 410 415

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp

			420					425					430		
Tyr	His	Ser 435	Gly	Tyr	Thr	Pro	Ser 440	Leu	Asp	Glu	Tyr	Leu 445	Asn	Ile	Ala
Lys	Ile 450	Ser	Val	Ala	Ser	Pro 455	Ala	Ile	Ile	Ser	Pro 460	Thr	Tyr	Phe	Thr
Phe 465	Ala	Asn	Ala	Ser	His 470	Asp	Thr	Ala	Val	Ile 475	Asp	Ser	Leu	Tyr	Gln 480
Tyr	His	qzA	Ile	Leu 485	Cys	Leu	Ala	Gly	Ile 490	Ile	Leu	Arg	Leu	Pro 495	Asp
Asp	Leu	Gly	Thr 500	Ser	Tyr	Phe .	Glu	Leu 505	Ala	Arg	Gly	Asp	Val 510	Pro	Lys
Thr	Ile	Gln 515	Cys	Tyr	Met	Lys	Glu 520	Thr	Asn	Ala	Ser	Glu 525	Glu	Glu	Ala
Val	Glu 530	His	Val	Lys	Phe	Leu 535	Ile	Arg	Glu	Ala	Trp 540	Lys	Asp	Met	Asn
Thr 545	Ala	Ile	Ala	Ala	Gly 550	Tyr	Pro	Phe	Pro	Asp 55 5	Gly	Met	Val	Ala	Gly 560
Ala	Ala	Asn	Ile	Gly 565	Arg	Val	Ala	Gln	Phe 570	Ile	Tyr	Leu	His	Gly 575	
			580			Ser	Lys	Thr 585	Tyr	Glu	His	Ile	Ala 590	Gly	Leu
Leu	Phe	Glu 595	Pro	Tyr	Ala										
<210 <211 <212 <213	> 20 > DN	24 A	cial	Seq	uenc	: e									٠
<220															
	CO	scri mput	ptio er-g	n of ener	Art ated	ific nuc	ial leic	Sequ aci	ence d se	: quen	ce				
<220 <221 <222 <223	> CD > (1 > co	1) mput	er-g	ener	ated hosp	nuc hate	leic syn	aci thas	d se	quen	ce e	ncod	ing		
<400															
gatc	acaa	aa a M	tg t et S 1	ct a er I	tc a le I	tt a le S	gc a er M 5	tg a et A	ac g sn V	ta t al S	er I	tc c le L 10	tt a eu S	gc a er L	ag 4: ys
CCa Pro	cta Leu . 15	aat Asn	tgc Cys	ctc Leu	cac His .	aac Asn 20	ttg Leu	gag Glu	agg : Arg :	aga Arg	cct Pro:	tca Ser	aaa Lys .	gcc Ala	ttg 9' Leu

ctt Leu 30	1 AGT	e cct	tgo Cys	act Thr	gca Ala 35	Pro	acc Thr	gct	egc Arg	cto Leu 40	Arg	gca Ala	tct Ser	tgc Cys	ser Ser 45	145
tca Ser	aaa : Lys	cta Leu	caa Gln	gaa Glu 50	Ala	cat His	caa Gln	atc	cga Arg 55	cga Arg	tct Ser	gga Gly	aac Asn	tac Tyr 60	Caa Gln	193
Pro	gcc Ala	ctt Leu	Trp 65	Asp	tcc Ser	aat Asn	tac Tyr	att Ile 70	cag Gln	tct Ser	ctc Leu	aat Asn	act Thr 75	cca Pro	tat Tyr	241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	ttg Leu	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	ctg Leu	caa Gln	gtg Val	289
agg Arg	ata Ile 95	ctg Leu	cta Leu	aag Lys	gaa Glu	aaa Lys 100	atg Met	gag Glu	cct Pro	gtt Val	caa Gln 105	caa Gln	ttg Leu	gag Glu	ttg Leu	337
att Ile 110	nıs	gac Asp	ttg Leu	aaa Lys	tat Tyr 115	ttg Leu	G ly ggg	ctc Leu	tcg Ser	gat Asp 120	ttt Phe	ttt Phe	caa Gln	gat Asp	gag Glu 125	385
att Ile	aag Lys	gag Glu	atc Ile	tta Leu 130	ggt Gly	gtt Val	ata Ile	tac Tyr	aat Asn 135	gag Glu	cac His	aaa Lys	tgc Cys	ttt Phe 140	cac His	433
aat Asn	aat Asn	gaa Glu	gta Val 145	gag Glu	aaa Lys	atg Met	gat Asp	ttg Leu 150	tat Tyr	ttc Phe	aca Thr	gct Ala	ctt Leu 155	gga Gly	ttc Phe	481
aga Arg	ctc Leu	ctc Leu 160	aga Arg	caa Gln	cat His	ggt Gly	ttt Phe 165	aat Asn	att Ile	tcc Ser	caa Gln	gat Asp 170	gta Val	ttt Phe	aat Asn	529
tgt Cys	ttc Phe 175	aag Lys	aac Asn	gag Glu	aag Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	agc Ser	ctt Leu	gct Ala	caa Gln	577
gat Asp 190	acg Thr	aag Lys	gga Gly	atg Met	tta Leu 195	caa Gln	ctg Leu	tat Tyr	gaa Glu	gcg Ala 200	tct Ser	ttc Phe	ctt Leu	ttg Leu	aga Arg 205	625
aaa Lys	ggt Gly	gaa Glu	gat Asp	aca Thr 210	ttg Leu	gag Glu	ctt Leu	Ala	aga Arg 215	gaa Glu	ttt Phe	gcc Ala	aca Thr	aaa Lys 220	tgt Cys	673
ctg Leu	cag Gln	aaa Lys	aaa Lys 225	ctt Leu	gat Asp	gaa Glu	GIA	ggt Gly 230	aat Asn	gaa Glu	att Ile	Asp	gag Glu 235	aat Asn	cta Leu	721
tta Leu	ttg Leu	tgg Trp 240	att Ile	ege Arg	cac His	tct Ser	ttg Leu 245	gat Asp	ctt Leu	cct Pro	ctc Leu	cac His 250	tgg Trp	agg Arg	att Ile	769
	agt Ser 255	gta Val	gag Glu	gca Ala	Arg	tgg Trp 260	ttc Phe	ata Ile	gat Asp	Ala	tat Tyr 265	gcg Ala .	aga Arg	agg Arg	cca Pro	817

ga As ₁ 27	y Me	g aat : Ası	t cca	ctt Leu	att Ile 275	Phe	gag Glu	ctt Leu	gco Ala	aaa Lys 280	Leu	aac Asn	tto Phe	aat Asn	att Ile 285	865
att Ile	caa Glr	gca Ala	a aca Thr	cat His 290	GIN	Caa Gln	gaa Glu	ctg Leu	Lys 295	: Asp	cto Leu	tcg Ser	agg Arg	tgg Trp 300	tgg Trp	913
agt Ser	aga Arg	tta Lev	tgc Cys 305	rne	cct Pro	gaa Glu	aag Lys	ctc Leu 310	Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	agg Arg	ctc Leu	961
gtt Val	gaa Glu	Ser 320	rne	ttt Phe	tgg Trp	gcg Ala	gtt Val 325	Gly	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
gga Gly	tat Tyr 335	GTD	aga Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Thr	att Ile	att Ile	gtt Val 345	tta Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	gat Asp	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt	aca Thr	cta Leu 360	gat Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac A sp	acg Thr	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	ctt Leu	1153
cct Pro	tat Tyr	tac Tyr	atg Met 385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 39 0	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
gat Asp	gca Ala	gca Ala 400	tat Tyr	gat Asp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caa Gln	1249
tat Tyr	ctc Leu 415	Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297
gca Ala 430	aag Lys	tgg Trp	tac Tyr	cac His	agc Ser 435	ggt Gly	tat Tyr	aca Thr	cca Pro	agc Ser 440	ctg Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
	atc Ile	Λια	БÅЗ	450	ser	Val	Ala	ser	Pro 455	Ala	Ile	Ile	Ser	Pro 460	Thr	1393
tat Tyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	tct Ser	cat His 470	gac Asp	aca Thr	gca Ala	gtc Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	caa Gln 480	tat Tyr	cat His	gac Asp	TTE	ctt Leu 485	tgc Cys	cta Leu	gca Ala	Gly	att Ile 490	att Ile	ttg Leu .	agg Arg	1489
ctt Leu	Pro 495	gac Asp	gat Asp	ctt Leu	GTA	aca Thr 500	tca Ser '	tat Tyr	ttt Phe	GLu	ctg Leu 505	gcg Ala	aga Arg	ggc Gly	gac Asp	1537

at																
Va. 51	g ccg L Pro	aaa Lys	aca Thr	atc Ile	Gln 515	Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585
ga Gl	g gag ı Glu	gcg Ala	gtg Val	gag Glu 530	His	gtg Val	aag Lys	ttt Phe	ctg Leu 535	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp 540	a ag Lys	1633
gat Ası	atg Met	aac Asn	acg Thr 545	gcc Ala	ata Ile	gca Ala	gcc Ala	ggt Gly 550	tat Tyr	ccg Pro	ttt Phe	ccg Pro	gat Asp 555	ggt Gly	atg Met	1681
gtç Val	gcg Ala	ggc Gly 560	gca Ala	gct Ala	aat Asn	att Ile	ggg Gly 56 5	cgc Arg	gtg Val	gcg Ala	cag Gln	ttt Phe 570	att Ile	tat Tyr	ctc Leu	1729
cac His	gga Gly 575	gat Asp	ej A aaa	ttt Phe	ggc Gly	gtg Val 580	caa Gln	cac His	tcg Ser	aaa Lys	acg Thr 585	tac Tyr	gag Glu	cat His	atc Ile	1777
gcc Ala 590	ggc	cta Leu	ctg Leu	ttc Phe	gag Glu 595	cct Pro	tat Tyr	gca Ala	tgaa	acaaa	atg g	gaga	ctg	et		1824
tga	tata	cat 1	taatt	tgg	ca ca	accaa	ıtaat	: tac	atot	tat	atat	atta	ıga a	aata	agtgt	1884
															tgtag	
															gctcg	
								- uu	.caac		aaac	agu	.ya ç	CCac	igeceg	
	tcga							, uu	·caac		aaac	aycc	.ya y	, ccae	igeeeg	2024
<21 <21 <21	tcgaq 0> 2° 1> 59 2> PI	gga a 7 8 RT	aaaa	laaaa	aa				· cauc		aaac	agu	ya y	, ccae	.gcccg	
<21 <21 <21 <21	0> 2° 1> 59 2> PI 3> A	gga a 98 RT Stifi	aaaa	laaaa	aa				···		aast	agu	ya <u>.</u>	, ccae	geeeg	
<21 <21 <21 <21	tcgaq 0> 2° 1> 59 2> PI	gga a 98 RT :tifi	aaaaa icial	aaaaa . Sec	ia Juenc	:e										
<21 <21 <21 <21 <40 Met 1	0> 2: 1> 5: 2> PI 3> A: 0> 2: Ser	gga a 98 RT tifi , Ile	icial Ile Asn 20	Ser Ser 5	uenc Met Glu	e Asn Arg	Val Arg	Ser Pro 25	Ile 10 Ser	Leu Lys	Ser Ala	Lys Leu	Pro Leu 30	Leu 15 Val	Asn Pro	
<pre><21 <21 <21 <40 Met Cys</pre>	0> 2° 1> 5° 2> PI 3> Ar 0> 2° Ser Leu	gga a 98 RT ctifi , Ile His Ala 35	icial Ile Asn 20 Pro	Ser 5 Leu	quenc Met Glu Ala	e Asn Arg Arg	Val Arg Leu 40	Ser Pro 25 Arg	Ile 10 Ser Ala	Leu Lys Ser	Ser Ala Cys	Lys Leu Ser 45	Pro Leu 30 Ser	Leu 15 Val Lys	Asn Pro Leu	
<pre><21 <21 <21 <40 Met Cys</pre>	0> 2: 1> 5: 2> PI 3> A: 0> 2: Ser	gga a 98 RT ctifi , Ile His Ala 35	icial Ile Asn 20 Pro	Ser 5 Leu	quenc Met Glu Ala	e Asn Arg Arg	Val Arg Leu 40	Ser Pro 25 Arg	Ile 10 Ser Ala	Leu Lys Ser	Ser Ala Cys	Lys Leu Ser 45	Pro Leu 30 Ser	Leu 15 Val Lys	Asn Pro Leu	
<pre><21 <21 <21 <40 Met 1 Cys Cys Gln</pre>	0> 2° 1> 5° 2> PI 3> Ar 0> 2° 5er Leu Thr	gga a 98 RT tifi , Ile His Ala 35	icial Ile Asn 20 Pro	Sec Ser 5 Leu Thr	na quenc Met Glu Ala Ile	e Asn Arg Arg	Val Arg Leu 40 Arg	Ser Pro 25 Arg Ser	Ile 10 Ser Ala Gly	Leu Lys Ser Asn	Ser Ala Cys Tyr 60	Lys Leu Ser 45 Gln	Pro Leu 30 Ser	Leu 15 Val Lys Ala	Asn Pro Leu Leu	
<pre><21 <21 <21 <40 Met 1 Cys Gln Trp 65 Arg</pre>	0> 2: 1> 5: 2> PI 3> AI 0> 2: Ser Leu Thr	gga a 98 RT tifi , Ile His Ala 35 Ala Ser Leu	icial Ile Asn 20 Pro His Asn Asn	Ser Ser 5 Leu Thr Gln Tyr Arg	Met Glu Ala Ile 70 Lys	Asn Arg Arg 55 Gln	Val Arg Leu 40 Arg Ser	Ser Pro 25 Arg Ser Leu	Ile 10 Ser Ala Gly Asn Ile 90	Leu Lys Ser Asn Thr 75 Leu	Ser Ala Cys Tyr 60 Pro Gln	Lys Leu Ser 45 Gln Tyr	Pro Leu 30 Ser Pro Thr	Leu 15 Val Lys Ala Glu Ile 95	Asn Pro Leu Leu Glu 80 Leu	

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 135 Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu L**e**u Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 295 Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 425

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 475

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys
500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 550 555

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 28

<211> 2024

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (11) .. (1804)

<223> computer-generated nucleic acid sequence encoding
 (+)-bornyl diphosphate synthase

<400> 28

gatcacaaaa atg tct atc att agc atg aac gta tcg atc ctt agc aag
Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys
1 5

CCa cta aat tgc ctc cac aac ttg gag agg aga cct tca aaa gcc ttg 97
Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu
15 20 25

ctt gtc cct tgc act gca ccc acc gct cgc ctc cgg gca tct tgc tcc 145 Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser

30					35					40					45	
tca Ser	aaa Lys	Cta Leu	caa Gln	gaa Glu 50	gct Ala	cat His	caa Gln	atc Ile	cga Arg 55	c ga Arg	tct Ser	gga Gly	aac Asn	tac Tyr 60	caa Gln	193
cct Pro	gcc	ctt Leu	tgg Trp 65	gat Asp	tcc Ser	aat Asn	tac Tyr	att Ile 70	cag Gln	tct Ser	ctc Leu	aat Asn	act Thr 75	cca Pro	tat Tyr	241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	ttg Leu	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	gtg Val	caa Gln	gtg Val	289
agg Arg	cta Leu 95	ctg Leu	cta Leu	aag Lys	gaa Glu	aaa Lys 100	atg Met	gag Glu	cct Pro	gtt Val	caa Gln 105	caa Gln	ttg Leu	gag Glu	ttg Leu	337
att Ile 110	cat His	gac Asp	ttg Leu	aaa Lys	tat Tyr 115	ttg Leu	Gly ggg	ctc Leu	tcg Ser	gat Asp 120	ttt Phe	ttt Phe	caa Gln	gat Asp	gag Glu 125	385
att Ile	a ag Lys	gag Glu	atc Ile	tta Leu 130	ggt Gly	gtt Val	ata Ile	tac Tyr	aat Asn 135	gag Glu	cac His	aaa Lys	tgc Cys	ttt Phe 140	cac His	433
aat Asn	aat Asn	gaa Glu	gta Val 145	gag Glu	aaa Lys	atg Met	gat Asp	ttg Leu 150	tat Tyr	ttc Phe	aca Thr	gct Ala	ctt Leu 155	gga Gly	ttc Phe	481
aga Arg	ctc Leu	ctc Leu 160	aga Arg	caa Gln	cat His	ggt Gly	ttt Phe 165	aat Asn	att Ile	tcc Ser	caa Gln	gat Asp 170	gta Val	ttt Phe	aat Asn	529
tgt Cys	ttc Phe 175	aag Lys	aac Asn	gag Glu	aag Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	agc Ser	ctt Leu	gct Ala	caa Gln	577
gat Asp 190	acg Thr	aag Lys	gga Gly	atg Met	tta Leu 195	caa Gln	ctg Leu	tat Tyr	gaa Glu	gcg Ala 200	tct Ser	ttc Phe	ctt Leu	ttg Leu	aga Arg 205	625
aaa Lys	ggt Gly	gaa Glu	gat Asp	aca Thr 210	ttg Leu	gag Glu	ctt Leu	gca Ala	aga Arg 215	gaa Glu	ttt Phe	gcc Ala	aca Thr	aaa Lys 220	tgt Cys	673
ctg Leu	cag Gln	Lys	aaa Lys 225	ctt Leu	gat Asp	gaa Glu	ggt Gly	ggt Gly 230	aat Asn	gaa Glu	att Ile	gat Asp	gag Glu 235	aat Asn	cta Leu	721
tta Leu	ttg Leu	tgg Trp 240	att Ile	cgc Arg	cac His	tct Ser	ttg Leu 245	gat Asp	ctt Leu	ect Pro	ctc Leu	cac His 250	tgg Trp	agg Arg	att Ile	769
caa Gln	agt Ser 255	gta Val	gag Glu	gca Ala	aga Arg	tgg Trp 260	ttc Phe	ata Ile	gat As p	gct Ala	tat Tyr 265	gcg Ala	aga Arg	agg Arg	cca Pro	817
gac Asp	atg Met	aat Asn	cca Pro	ctt Leu	att Ile	ttc Phe	gag Glu	ctt Leu	gcc Ala	aaa Lys	ctc Leu	aac Asn	ttc Phe	aat Asn	att Ile	865

270)				275					280)				285	
att Ile	caa Gln	gca Ala	aca Thr	Cat His 290	Caa Gln	caa Gln	gaa Glu	ctg Leu	aaa Lys 295	Asp	ctc Leu	tcg Ser	agg Arg	tgg Trp 300	Trp	913 *
agt Ser	aga Arg	tta Leu	tgc Cys 305	ttc Phe	Pro	gaa Glu	aag Lys	ctc Leu 310	Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	agg A rg	ctc Leu	961
gtt Val	gaa Glu	tcc Ser 320	Phe	ttt Phe	tgg Trp	gcg Ala	gtt Val 325	Gly	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
gga Gly	tat Tyr 335	cag Gln	aga Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Thr	att Ile	att Ile	gtt Val 345	tta Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	gat Asp	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt Gly	aca Thr	cta Leu 360	gat Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac Asp	acg Thr	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	ctt Leu	1153
ect Pro	tat Tyr	tac Tyr	atg Met 385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 390	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
gat Asp	gca Ala	gca Ala 400	tat Tyr	gat Asp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caa Gln	1249
tat Tyr	ctc Leu 415	cgg Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297
gca Ala 430	aag Lys	tgg Trp	tac Tyr	cac His	agc Ser 435	ggt Gly	tat Tyr	aca Thr	cca Pro	agc Ser 440	ctg Leu	gat Asp	gaa Glu	tat Tyr	ctc Leu 445	1345
aac Asn	atc Ile	gcc Ala	aag Lys	att Ile 450	tca Ser	gtg Val	gcg Ala	tct Ser	cct Pro 455	gca Ala	ata Ile	ata Ile	tcc Ser	cca Pro 460	acc Thr	1393
tat Tyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	Ala	Ser	cat His 470	qzA	aca Thr	gca Ala	Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	caa Gln 480	tat Tyr	cat His	gac Asp	ata Ile	ctt Leu 485	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile 490	att Ile	ttg Leu	agg Arg	1489
~~ u	ecc Pro 495	gac Asp	gat Asp	ctt Leu	ej A aaa	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	gcg Ala	aga Arg	ggc Gly	gac ' Asp	1537
gtg Val	ccg Pro	aaa Lys	aca Thr	atc Ile	cag Gln	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu	aca Thr	aat Asn	gct Ala	agt Ser	gag Glu	1585

510 515 520 525 gag gag gcg gtg gag cac gtg aag ttt ctg ata agg gag gcg tgg aag 1633 Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys 530 gat atg aac acg gcc ata gca gcc ggt tat ccg ttt ccg gat ggt atg 1681 Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met 545 . 550 555 gtg gcg ggc gca gct aat att ggg cgc gtg gcg cag ttt att tat ctc 1729 Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu 565 cac gga gat ggg ttt ggc gtg caa cac tcg aaa acg tac gag cat atc His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile 580 gcc ggc cta ctg ttc gag cct tat gca tgaacaaatg ggagactgct 1824 Ala Gly Leu Leu Phe Glu Pro Tyr Ala 590 tgatatatat taatttggca caccaataat tgcatgttat atatgttgga aaataagtgt 1884 ctggttgaga tgtcatgtgg tgtattatct aaataattca aggttgcctt gtttatgtag 1944 ccggtggtgc aactacctcc cattcaaatc aattaaatct aaacagtcga gtcaagctcg 2004 agctcgagga aaaaaaaaa <210> 29 <211> 598 <212> PRT <213> Artificial Sequence <400> 29 Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Leu Leu Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 120

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 440

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asm Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu
580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 30

<211> 2024

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (11)..(1804)

<223> computer-generated nucleic acid sequence encoding
 (+)-bornyl diphosphate synthase

<400> 30

gatcacaaaa atg tct atc att agc atg aac gta tcg atc ctt agc aag
Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys
1 5

CCa Cta aat tgc Ctc Cac aac ttg gag agg aga CCt tca aaa gCC ttg

Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu

15
20
25

ctt gtc cct tgc act gca ccc acc gct cgc ctc cgg gca tct tgc tcc 145
Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser
30 45

tca Ser	aaa Lys	cta Leu	caa Gln	gaa Glu 50	Ala	cat His	caa Gln	atc Ile	cga Arg 55	Arg	tct Ser	gga Gly	aac Asn	tac Tyr 60	Gln	193
ect Pro	gcc Ala	ctt Leu	tgg Trp 65	Asp	tcc Ser	aat Asn	tac Tyr	att Ile 70	cag Gln	tct Ser	ctc Leu	aat Asn	act Thr 75	cca Pro	tat Tyr	241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	ttg Leu	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	gtg Val	caa Gln	gtg Val	289
Arg	95	Val	cta ·Leu	Lys	Glu	Lys 100	Met	Glu	Pro	Val	Gln 105	Gln	Leu	Glu	Leu	337
110	His	Asp	ttg Leu	Lys	Tyr 115	Leu	Gly	Leu	Ser	Asp 120	Phe	Phe	Gln	Asp	Glu 125	385
116	гÀз	Glu	atc Ile	Leu 130	Gly	Val	Ile	Tyr	Asn 135	Glu	His	Lys	Суз	Phe 140	His	433
Asn	Asn	GIu	gta Val 145	Glu	Lys	Met	Asp	Leu 150	Tyr	Phe	Thr	Ala	Leu 155	Gly	Phe	481
Arg	ren	160	aga Arg	Gln	His	Gly	Phe 165	Asn	Ile	Ser	Gln	Asp 170	Val	Phe	Asn	529
Суз	175	Lys	aac Asn	Glu	Lys	Gly 180	Ile	Asp	Phe	Lys	Ala 185	Ser	Leu	Ala	Gln	577
190	Thr	Lys	gga Gly	Met	Leu 195	Gln	Leu	Tyr	Glu	Ala 200	Ser	Phe	Leu	Leu	Arg 205	625
rĀs	СТÀ	Glu	gat Asp	Thr 210	Leu	Glu	Leu	Ala	Arg 215	Glu	Phe	Ala	Thr	Lys 220	Суз	673
red	GIN	Lys	aaa Lys 225	Leu	Asp	Glu	Gly	Gly 230	Asn	Glu	Ile	Asp	Glu 235	Asn	Leu	721
beu	rea	240	att Ile	Arg	His	Ser	Leu 245	Asp	Leu	Pro	Leu	His 250	Trp	Arg	Ile	769
GIN	255	Val	gag Glu	Ala	Arg	Trp 260	Phe	Ile	Asp	Ala	Tyr 265	Ala	Arg	Arg	Pro	817
gac Asp 270	atg Met	aat Asn	cca Pro	Leu	att Ile 275	ttc Phe	gag Glu	ctt Leu	Ala	aaa Lys 280	ctc Leu	aac Asn	ttc Phe	aat Asn	att Ile 285	865

at Il	t ca e Gl	a gc n Al	a aca	a cat r His 290	, GID	Cas Gln	gaa Glu	cto Lev	7 aas Lys 295	s Asi	t ct	c tc	g ag	g tg g Trj 30	g tgg p Trp	913
agi Sei	t aga r Arq	a tt g Le	a tgo u Cys 305	> Ene	cct Pro	gaa Glu	aag Lys	ctc Leu 310	cca Pro	. +++	gte Val	g ago	g gat g Asi 315	age Are	g ctc g Leu	961
gt(Va)	t gaa L Glu	s to 1 Se: 32	r sue	ttt Phe	tgg Trp	gcg Ala	gtt Val 325	Gly	atg Met	r ttt : Phe	gaç Glu	g cca Pro 330) His	Caa Glr	cat His	1009
G1 y	tat Tyr 335	. GI:	g aga n Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Thr	att Ile	att Ile	gtt Val 345	. Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	աբ	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt Gly	aca Thr	cta Leu 360	Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac Asp	acg Thr	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	Leu	1153
cct Pro	tat Tyr	tac Tyr	: atg : Met :385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 390	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
gat Asp	gca Ala	gca Ala 400	TAL	gat Asp	att Ile	ctc Leu	aaa Lys 405	gaa Glu	cat His	ggt Gly	ttc Phe	ttt Phe 410	tgt Cys	ctc Leu	caa Gln	1249
tat Tyr	ctc Leu 415	cgg Arg	aaa Lys	tcg Ser	gtg Val	gta Val 420	gat Asp	ttg Leu	gtt Val	gaa Glu	gca Ala 425	tat Tyr	ttt Phe	cac His	gag Glu	1297
430	Lys	TIP	TYE	nis	agc Ser 435	GLY	Tyr	Thr	Pro	Ser 440	Leu	Asp	Glu	Tyr	Leu 445	1345
		, LL	Dys	450	tca Ser	val	ATA	Ser	Pro 455	Ala	Ile	Ile	Ser	Pro 460	Thr	1393 [.]
-,-		****	465	ATG	aac Asn	ALA	Ser	470	Asp	Thr	Ala	Val	11e 475	Asp	Ser	1441
	•	480	-,-		gac Asp	116	485	Cys	ren	Ala	СТĀ	11e 490	Ile	Leu	Arg	1489
	495	- - -				500	Jer .	ıyr .	rne	GIU	Leu 505	Ala	Arg	Gly	Asp	1537
gtg Val 510	ccg Pro	aaa Lys	aca Thr	116	cag (Gln (515	tgc (Cys (tac a	atg : Met :	rys	gaa Glu ' 520	aca Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585

gag gag gcg gtg gag cac gtg aag ttt ctg ata agg gag gcg tgg aag 10 Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys 530 535 540	633
gat atg aac acg gcc ata gca gcc ggt tat ccg ttt ccg gat ggt atg 16 Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met 545 550 555	681
gtg gcg ggc gca gct aat att ggg cgc gtg gcg cag ttt att tat ctc 17 Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu 560 565 570	729
cac gga gat ggg ttt ggc gtg caa cac tcg aaa acg tac gag cat atc 17 His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile 575 580 585	777
gcc ggc cta ctg ttc gag cct tat gca tgaacaaatg ggagactgct 18 Ala Gly Leu Leu Phe Glu Pro Tyr Ala 590 595	824
tgatatatat taatttggca caccaataat tgcatgttat atatgttgga aaataagtgt 18	384
ctggttgaga tgtcatgtgg tgtattatct aaataattca aggttgcctt gtttatgtag 19	34 4
eeggtggtgc aactacetee catteaaate aattaaatet aaacagtega gteaageteg 20	004
agctcgagga aaaaaaaaa 20	024
<210> 31	

<211> 598

<212> PRT

<213> Artificial Sequence

<400> 31

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 10 15

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40 45

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Val 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu

130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155 160

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 215 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 225 230 235

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 245 250 250

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 265 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 290 295 300

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 310 315 320

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 340 345 350

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355 360 365

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 370 375 380

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg 405 410 415

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 420 425 430

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr

	450					455					460					
Phe 465	Ala	Asn	Ala	Ser	His 470	Asp	Thr	Ala	Val	Ile 475	Asp	Ser	Leu	Tyr	Gln 480	
Tyr	His	Asp	Ile	Leu 485	Cys	Leu	Ala	Gly	Ile 490	Ile	Leu	Arg	Leu	Pro 495	Asp	
Asp	Leu	Gly	Thr 500	Ser	Tyr	Phe	Glu	Leu 505	Ala	Arg	Gly	Asp	Val 510	Pro	Lys	
Thr	Ile	Gln 515	Cys	Tyr	Met	Lys	Glu 520	Thr	Asn	Ala	Ser	Glu 525	Glu	Gl u	Ala	
Val	Glu 530	His	Val	Lys.	Phe	Leu 535	Ile	Arg	Glu	Ala	Trp 540	Lys	Asp	Met	Asn	
Thr 545	Ala	Ile	Ala	Ala	Gly 550	Tyr	Pro	Phe	Pro	Asp 555	Gly	Met	Val	Ala	Gly 560	
Ala	Ala	Asn	Ile	Gly 565	Arg	Val	Ala	Gln	Phe 570	Ile	Tyr	Leu	His	Gly 575	Asp	
Gly	Phe	Gly	Val 580	Gln	His	Ser	Lys	Thr 585	Tyr	Glu	His	Ile	Ala 590	Gly	Leu	
Leu	Phe	Glu 595	Pro	Tyr	Ala											
<21:	0> 32 L> 20 2> DR 3> A1	024 VA	lcia]	L Sec	juenc	:e										
<223	3> De	escri	ptic er-c	enei	E Art	ific i nuc	ial :leid	Sequ aci	ence	: :quen	ce					
<222	l> CI 2> (1 3> cc	11)	er-g	ene	ateo	i nuc hate	:leid : syr	aci athas	.d se	quen	ice e	ncoc	ling			
)> 32 :acaa	iaa a	tg t let S	ct a Ser 1	itc a [le]	itt a :le S	igc a Ser M	itg a Met A	ac g sn V	rta t Val S	cg a er I	tc c le I 10	tt a eu S	gc a er L	ag ys	49
cca Pro	cta Leu 15	aat Asn	tgc Cys	ctc Leu	cac His	aac Asn 20	ttg Leu	gag Glu	agg Arg	aga Arg	ect Pro 25	tca Ser	aaa Lys	gcc Ala	ttg Leu	97
ctt Leu 30	gtc Val	cct Pro	tgc Cys	act Thr	gca Ala 35	ccc Pro	acc Thr	gct Ala	cgc Arg	ctc Leu 40	egg Arg	gca Ala	tct Ser	tgc Cys	tec Ser 45	145
tca Ser	aaa Lys	cta Leu	caa Gln	gaa Glu	gct Ala	cat His	caa Gln	atc Ile	cga Arg	cga Arg	tct Ser	gga Gly	aac Asn	tac Tyr	caa Gln	193

193

				50					55					60		
					tcc Ser											241
acg Thr	gag Glu	gag Glu 80	agg Arg	cac His	ttg Leu	gat Asp	aga Arg 85	aaa Lys	gca Ala	gag Glu	ctg Leu	att Ile 90	gtg Val	caa Gln	gtg Val	289
agg Arg	ata Ile 95	ctg Leu	ata Ile	aag Lys	gaa Glu	aaa Lys 100	a tg Met	gag Glu	cct Pro	gtt Val	caa Gln 105	caa Gln	ttg Leu	gag Glu	ttg Leu	337
					tat Tyr 115											385
					ggt Gly											433
					aaa Lys											481
					cat His											529
tgt Cys	ttc Phe 175	aag Lys	aac Asn	gag Glu	aag Lys	ggt Gly 180	att Ile	gat Asp	ttc Phe	aag Lys	gca Ala 185	agc Ser	ctt Leu	gct Ala	caa Gln	577
					tta Leu 195											625
aaa Lys	ggt Gly	gaa Glu	gat Asp	aca Thr 210	ttg Leu	gag Glu	ctt Leu	gca Ala	aga Arg 215	gaa Glu	ttt Phe	gcc Ala	aca Thr	aaa Lys 220	tgt Cys	673
ctg Leu	cag Gln	aaa Lys	aaa Lys 225	ctt Leu	gat Asp	gaa Glu	ggt Gly	ggt Gly 230	aat Asn	gaa Glu	att Ile	gat Asp	gag Glu 235	aat Asn	cta Leu	721
tta Leu	ttg Leu	tgg Trp 240	Ile	ege Arg	cac His	Ser	Leu	Asp	ctt Leu	Pro	ctc Leu	cac His 250	Trp	agg Arg	att Ile	769
caa Gln	agt Ser 255	gta Val	gag Glu	gca Ala	aga Arg	tgg Trp 260	ttc Phe	ata Ile	gat Asp	gct Ala	tat Tyr 265	gcg Ala	aga Arg	agg Arg	cca Pro	817
gac Asp 270	atg Met	aat Asn	cca Pro	ctt Leu	att Ile 275	ttc Phe	gag Glu	ctt Leu	gcc Ala	aaa Lys 280	ctc Leu	aac Asn	ttc Phe	aat Asn	att Ile 285	865
att Ile	caa Gln	gca Ala	aca Thr	cat His	caa Gln	caa Gln	gaa Glu	ctg Leu	aaa Lys	gat As p	ctc Leu	tcg Ser	agg Arg	tgg Trp	tgg Trp	913

				290					295					300		
agt Ser	aga Arg	tta Leu	tgc Cys 305	ttc Phe	Pro	gaa Glu	aag Lys	ctc Leu 310	cca Pro	ttt Phe	gtg Val	agg Arg	gat Asp 315	agg Arg	ctc Leu	961
gtt Val	gaa Glu	Ser 320	ttc Phe	ttt Phe	tgg Trp	g¢g Ala	gtt Val 325	GJA	atg Met	ttt Phe	gag Glu	cca Pro 330	cat His	caa Gln	cat His	1009
gga Gly	tat Tyr 335	cag Gln	aga Arg	aaa Lys	atg Met	gcc Ala 340	gcc Ala	aca Thr	att Ile	att Ile	gtt Val 345	tta Leu	gca Ala	aca Thr	gtt Val	1057
ata Ile 350	gat Asp	gat Asp	att Ile	tac Tyr	gat Asp 355	gtg Val	tat Tyr	ggt Gly	aca Thr	cta Leu 36 0	gat Asp	gaa Glu	cta Leu	gaa Glu	cta Leu 365	1105
ttt Phe	aca Thr	gac Asp	acg Thr	ttt Phe 370	aag Lys	aga Arg	tgg Trp	gat Asp	act Thr 375	gaa Glu	tca Ser	ata Ile	acc Thr	cga Arg 380	ctt Leu	1153
cct Pro	tat Tyr	tac Tyr	atg Met 385	caa Gln	tta Leu	tgt Cys	tat Tyr	tgg Trp 390	ggt Gly	gtc Val	cac His	aac Asn	tat Tyr 395	att Ile	tcc Ser	1201
Asp	Ala	Ala 400	tat Tyr	Asp	Ile	Leu	Lys 405	Glu	His	Gly	Phe	Phe 410	Cys	Leu	Gln	1249
туг	115	Arg	aaa Lys	Ser	Val	Val 420	Asp	Leu	Val	Glu	Ala 425	Tyr	Phe	His	Glu	1297
430	гÀЗ	Trp	tac Tyr	His	Ser 435	Gly	Tyr	Thr	Pro	Ser 440	Leu	Asp	Glu	Tyr	Leu 445	1345
ASN	TTE	Ala	aag Lys	11e 450	Ser	Val	Ala	Ser	Pro 455	Ala	Ile	Ile	Ser	Pro 460	Thr	1393
tat Tyr	ttc Phe	aca Thr	ttc Phe 465	gca Ala	aac Asn	gcg Ala	tct Ser	cat His 470	gac Asp	aca Thr	gca Ala	gtc Val	atc Ile 475	gac Asp	agc Ser	1441
ttg Leu	tac Tyr	Caa Gln 480	tat Tyr	cat His	gac Asp	ata Ile	ctt Leu 485	tgc Cys	cta Leu	gca Ala	gga Gly	att Ile 490	att Ile	ttg Leu	agg Arg	1489
ctt Leu	Pro 495	gac Asp	gat Asp	ctt Leu	G] À aaa	aca Thr 500	tca Ser	tat Tyr	ttt Phe	gag Glu	ctg Leu 505	gcg Ala	aga Arg	ggc Gly	gac As p	1537
gtg Val 510	ccg Pro	a aa Lys	aca Thr	atc Ile	cag Gln 515	tgc Cys	tac Tyr	atg Met	aag Lys	gaa Glu 520	aca Thr	aat Asn	gct Ala	agt Ser	gag Glu 525	1585
gag Glu	gag Glu	gcg Ala	gtg Val	gag Glu	cac His	gtg Val	aag Lys	ttt Phe	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala	tgg Trp	aag Lys	1633

530 535 540 gat atg aac acg gcc ata gca gcc ggt tat ccg ttt ccg gat ggt atg Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met 1681 545 550 gtg gcg ggc gca gct aat att ggg cgc gtg gcg cag ttt att tat ctc 1729 Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu cac gga gat ggg ttt ggc gtg caa cac tcg aaa acg tac gag cat atc His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile. 580 gcc ggc cta ctg ttc gag cct tat gca tgaacaaatg ggagactgct 1824 Ala Gly Leu Leu Phe Glu Pro Tyr Ala tgatatatat taatttggca caccaataat tgcatgttat atatgttgga aaataagtgt 1884 ctggttgaga tgtcatgtgg tgtattatct aaataattca aggttgcctt gtttatgtag 1944 ccggtggtgc aactacctcc cattcaaatc aattaaatct aaacagtcga gtcaagctcg 2004 agctcgagga aaaaaaaaa 2024 <210> 33 <211> 598 <212> PRT <213> Artificial Sequence <400> 33 Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu Gln Glu Ala Kis Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu Ile Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 105 Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 135

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 155 Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 235 Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400 Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 550 555

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 34

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(598)

<223> Computer-generated (+)-bornyl diphosphate synthase
protein variant

<400> 34

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 15

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 45

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Ile 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu
65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu

95

90

Leu	Lys	Glu	Lys 100	Met	Glu	Pro	Val	Gln 105	Gln	Leu	Glu	Leu	Ile 110	His	Asp
Leu	Lys	Tyr 115	Leu	Gly	Leu	Ser	Asp 120	Phe	Phe	Gln	Asp	Glu 125	Ile	Lys	Glu
Ile	Leu 130	Gly	Val	Ile	Tyr	Asn 135	Glu	His	Lys	Cys	Phe 140	His	Asn	Asn	Glu
Val 145	Glu	Lys	Met	Asp	Leu 150	Tyr	Phe	Thr	Ala	Leu 155	Gly	Phe	Arg	Leu	Leu 160
Arg	Gln	His	Gly	Phe 165	Asn	Ile	Ser	Gln	Asp 170	VáJ	Phe	Asn	Суз	Phe 175	Lys
Asn	Glu	Lys	Gly 180	Ile	Asp	Phe	Lys	Ala 185	Ser	Leu	Ala	Gln	Asp 190	Thr	Lys
Gly	Met	Leu 195	Gln	Leu	Tyr	Glu	Ala 200	Ser	Phe	Leu	Leu	Arg 205	Lys	Gly	Glu
Asp	Thr 210	Leu	Glu	Leu	Ala	Arg 215	Glu	Phe	Ala	Thr	Lys 220	Cys	Leu	Gln	Lys
Lys 225	Leu	Asp	Glu	Gly	Gly 230	Asn	Glu	Ile	Asp	Glu 235	Asn	Leu	Leu	Leu	Trp 240
Ile	Arg	His	Ser	Leu 245	Asp	Leu	Pro	Leu	His 250	Trp	Arg	Ile	Gln	Ser 2 5 5	Val
Glu	Ala	Arg	Trp 260	Phe	Ile	Asp	Ala	Tyr 265	Ala	Arg	Arg	Pro	Asp 270	Met	Asn
Pro	Leu	11e 275	Phe	Glu	Lėu	Ala	Lys 280	Leu	Asn	Phe	Asn	11e 285	Ile	Gln	Ala
Thr	His 290	Gln	Gln	Glu	Leu	Lys 295	Asp	Leu	Ser	Arg	Trp 300	Trp	Ser	Arg	Leu

85

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 350

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 370

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 305 310 315

405 410 415

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 420 425 430

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln
465 470 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585

Leu Phe Glu Pro Tyr Ala 595

<210> 35

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(598)

<223> Computer-generated (+)-bornyl diphosphate synthase sequence variant

<400> 35

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 10

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60 Trp Asp Ser Asn Tyr Leu Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80 Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95 Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125 Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335 Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 440 Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 455 Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 Leu Phe Glu Pro Tyr Ala 595

<210> 36

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(598)

<223> Computer-generated (+)-bornyl diphosphate synthase protein sequence variant

<400> 36

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 10

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Ile Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 225 230 235

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 245 250 250

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 305 310 315 320 Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Mot Gin Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala Lys Ilc Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 520 Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 535 Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu. 585 Leu Phe Glu Pro Tyr Ala

<210> 37 <211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated sequence variant

<220>

<221> VARIANT

<222> (1)..(598)

<223> Computer-generated (+)-bornyl diphosphate synthase protein variant

<400> 37

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 15

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Val Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 135 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 215 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp

225					230					235					240
Ile	Arg	His	Ser	Leu 245	Asp	Leu	Pro	Leu	His 250	Trp	Arg	Ile	Gln	Ser 255	Val
Glu	Ala	Arg	Trp 260	Phe	Ile	Asp	Ala	Tyr 265	Ala	Arg	Arg	Pro	Азр 270	Met	Asn
Pro	Leu	Ile 275	Phe	Glu	Leu	Ala	Lys 280	Leu	Asn	Phe	Asn	11e 285	Ile	Gln	Ala
Thr	His 290	Gln	Gln	Glu	Leu	Lys 295	Asp	Leu	Ser	Arg	Trp 300	Trp	Ser	Arg	Leu
Cys 305	Phe	Pro	Glu	Lys	Leu 310	Pro	Phe	Val	Arg	Asp 315	Arg	Leu	Val	Glu	Ser 320
Phe	Phe	Trp	Ala	Val 325	Gly	Met	Phe	Glu	Pro 330	His	Gln	His	Gly	Tyr 335	Gln
Arg	Lys	Met	Ala 340	Ala	Thr	Ile	Ile	Val 345	Leu	Ala	Thr	Val	Ile 350	Asp	Ąsp
Ile	Tyr	Asp 355	Val	Tyr	Gly	Thr	Leu 360	Asp	Glu	Leu	Glu	Leu 365	Phe	Thr	Asp
•	Phe 370					375					380				
385	Gln				390					395					400
	Asp			405					410					415	
	Ser		420					425					430		
	His	435					440					445			
	11e 450					455					460				
465	Ala				470					475					480
	His			485					490					495	
qzA	Leu	Gly	Thr 500	Ser	Tyr	Phe	Glu	Leu 505	Ala	Arg	Gly	Asp	Val 510	Pro	Lys
Thr	Ile	Gln 515	Cys	Tyr	Met	Lys	Glu 520	Thr	Asn	Ala	Ser	Glu 52 5	Glu	Glu	Ala
Val	Glu 530	His	Val	Lys	Phe	Leu 535	Ile	Arg	Glu	Ala	Trp 540	Lys	Азр	Met	Asn
Thr	Ala	Ile	Ala	Ala	Gly	Tyr	Pro	Phe	Pro	Asp	Gly	Met	Val	Ala	Gly

545 550 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 38 <211> 598

<211> 598 <212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(598)

<223> Computer-generated (+)-bornyl diphosphate synthase
protein sequence variant

<400> 38

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 10

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40 45

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75

Arg His Leu Asp Arg Lys Ala Glu Val Ile Val Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130 140

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155 160

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 280 Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 360 Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr 375 Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 390 395 Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 440 Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 470 Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585

Leu Phe Glu Pro Tyr Ala 595

<210> 39

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated sequence variant

<220>

<221> VARIANT

<222> (1)..(598)

<223> computer-generated (+)-bornyl diphosphate synthase
protein sequence variant

<400> 39

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn 1 5 15

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro 20 25 30

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40 45

Glm Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu 50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Leu Gln Val Arg Ile Leu 85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105 110

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 235 Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 425

.

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 465 470 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 590

Leu Phe Glu Pro Tyr Ala 595

<210> 40

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated sequence variant

<220>

<221> VARIANT

<222> (1)..(598)

<223> computer-generated (+)-bornyl diphosphate synthase protein sequence variant

<400> 40

Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn

Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro

Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu 35 40

Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu

50 55 60

Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu 65 70 75 80

Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Leu Leu
85 90 95

Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp 100 105

Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu 115 120 125

Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu 130
135

Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu 145 150 155 160

Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys 165 170 175

Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys 180 185 190

Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu 195 200 205

Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 210 220

Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp 225 230 235 240

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val 245 250 255

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn 260 270

Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala 275 280 285

Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 290 295 300

Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser 305 310 315 320

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln 325 330 335

Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp 340 345

Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp 355 360 365

Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr

370 375 .380

Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala 385 390 395 400

Tyr Asp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg 405 410 415

Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp
420 425 430

Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala 435 440 445

Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr 450 455 460

Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln 475 480

Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp 485 490 495

Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys 500 505 510

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly 545 550 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 585 590

Leu Phe Glu Pro Tyr Ala 595

<210> 41

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(598)

<223> computer-generated (+)-bornyl diphosphate synthase
protein sequence variant

<400> 41 Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Val Leu Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 215 Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu 295 Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser

. . ..

330

Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln

Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala 515 520 525

Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn 530 535 540

Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly
545 550 555 560

Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp 565 570 575

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 580 595

Leu Phe Glu Pro Tyr Ala 595

<210> 42

<211> 598

<212> PRT

<213> Artificial Sequence

<220> <223> Description of Artificial Sequence: computer-generated protein sequence <220> <221> VARIANT <222> (1)..(598) <223> computer-generated (+)-bornyl diphosphate synthase protein sequence variant <400> 42 Met Ser Ile Ile Ser Met Asn Val Ser Ile Leu Ser Lys Pro Leu Asn Cys Leu His Asn Leu Glu Arg Arg Pro Ser Lys Ala Leu Leu Val Pro Cys Thr Ala Pro Thr Ala Arg Leu Arg Ala Ser Cys Ser Ser Lys Leu Gln Glu Ala His Gln Ile Arg Arg Ser Gly Asn Tyr Gln Pro Ala Leu Trp Asp Ser Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Thr Glu Glu Arg His Leu Asp Arg Lys Ala Glu Leu Ile Val Gln Val Arg Ile Leu Ile Lys Glu Lys Met Glu Pro Val Gln Gln Leu Glu Leu Ile His Asp Leu Lys Tyr Leu Gly Leu Ser Asp Phe Phe Gln Asp Glu Ile Lys Glu Ile Leu Gly Val Ile Tyr Asn Glu His Lys Cys Phe His Asn Asn Glu Val Glu Lys Met Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Leu Leu Arg Gln His Gly Phe Asn Ile Ser Gln Asp Val Phe Asn Cys Phe Lys Asn Glu Lys Gly Ile Asp Phe Lys Ala Ser Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Lys Gly Glu Asp Thr Leu Glu Leu Ala Arg Glu Phe Ala Thr Lys Cys Leu Gln Lys 215 Lys Leu Asp Glu Gly Gly Asn Glu Ile Asp Glu Asn Leu Leu Trp

Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Ile Gln Ser Val

Glu Ala Arg Trp Phe Ile Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Glu Leu Ala Lys Leu Asn Phe Asn Ile Ile Gln Ala Thr His Gln Gln Glu Leu Lys Asp Leu Ser Arg Trp Trp Ser Arg Leu Cys Phe Pro Glu Lys Leu Pro Phe Val Arg Asp Arg Leu Val Glu Ser Phe Phe Trp Ala Val Gly Met Phe Glu Pro His Gln His Gly Tyr Gln Arg Lys Met Ala Ala Thr Ile Ile Val Leu Ala Thr Val Ile Asp Asp Ile Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Thr Phe Lys Arg Trp Asp Thr Glu Ser Ile Thr Arg Leu Pro Tyr Tyr Met Gln Leu Cys Tyr Trp Gly Val His Asn Tyr Ile Ser Asp Ala Ala **Tyr A**sp Ile Leu Lys Glu His Gly Phe Phe Cys Leu Gln Tyr Leu Arg Lys Ser Val Val Asp Leu Val Glu Ala Tyr Phe His Glu Ala Lys Trp 425 Tyr His Ser Gly Tyr Thr Pro Ser Leu Asp Glu Tyr Leu Asn Ile Ala Lys Ile Ser Val Ala Ser Pro Ala Ile Ile Ser Pro Thr Tyr Phe Thr Phe Ala Asn Ala Ser His Asp Thr Ala Val Ile Asp Ser Leu Tyr Gln Tyr His Asp Ile Leu Cys Leu Ala Gly Ile Ile Leu Arg Leu Pro Asp Asp Leu Gly Thr Ser Tyr Phe Glu Leu Ala Arg Gly Asp Val Pro Lys Thr Ile Gln Cys Tyr Met Lys Glu Thr Asn Ala Ser Glu Glu Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Asp Met Asn Thr Ala Ile Ala Ala Gly Tyr Pro Phe Pro Asp Gly Met Val Ala Gly Ala Ala Asn Ile Gly Arg Val Ala Gln Phe Ile Tyr Leu His Gly Asp

Gly Phe Gly Val Gln His Ser Lys Thr Tyr Glu His Ile Ala Gly Leu 585

Leu Phe Glu Pro Tyr Ala 595

<210> 43 <211> 1912 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (26)..(1792) <223> computer-generated nucleic acid sequence encoding (+)-sabinene synthase

<400> 43

agcaatatta caactaacaa taaaa atg tct tcc att agc ata aac ata gct Met Ser Ser Ile Ser Ile Asn Ile Ala

atg cca ctg aat tee ete cae aae ttt gag agg aaa eet tea aaa gea Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala

tgg tot acc tot tgc act gca ccc gca gct cgc ctc cgg gca tot tcc Trp Ser Thr Ser Cys Thr Ala Pro Ala Ala Arg Leu Arg Ala Ser Ser

tcc tta caa caa gaa aaa cct cac caa atc cga cgc tct ggg gat tac 196 Ser Leu Gln Gln Glu Lys Pro His Gln Ile Arg Arg Ser Gly Asp Tyr

caa ccc tct att tgg gat ttc aat tac ata cag tct ctc aac act ccg 244 Gln Pro Ser Ile Trp Asp Phe Asn Tyr Ile Gln Ser Leu Asn Thr Pro 60 65

tat aag gag cag aga cac ttt aat agg caa gca gag ttg att atg caa 292 Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ala Glu Leu Ile Met Gln 80

gtg agg atg ttg ctc aag gta aag atg gag gca att caa cag ttg gag 340 Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu 100

ttg att gat gac ttg caa tac ctg gga ctg tct tat ttc ttt caa gat 388 Leu Ile Asp Asp Leu Gln Tyr Leu Gly Leu Ser Tyr Phe Phe Gln Asp 110 115

gag att aaa caa atc tta agt tot ata cac aat gag ccc aga tat tto 436 Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe 125 130

cac aat aat gat ttg tat ttc aca gct ctt gga ttc aga atc ctc aga

His	Asn	Asn 140	Asp	Leu	Tyr	Phe	Thr 145		Leu	Gly	Phe	Arg 150		Leu	Arg	
caa Gln	His 155	GIY	ttt Phe	eat Asn	gtt Val	Ser 160	Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile	532
gag Glu 170	rys	tgc Cys	agt Ser	gat Asp	ttc Phe 175	Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	Gln	gat Asp	acg Thr	aag Lys	gga Gly 185	580
atg Met	tta Leu	caa Gln	ctt Leu	tat Tyr 190	gaa Glu	gca Ala	tct Ser	ttc Phe	ctt Leu 195	Leu	aga Arg	gaa Glu	ggt Gly	gaa Glu 200	gat As p	628
aca Thr	ttg Leu	gag Glu	cta Leu 205	gca Ala	aga Arg	cga Arg	ttt Phe	tcc Ser 210	acc Thr	aga Arg	tct Ser	cta Leu	cga Arg 215	gaa Glu	aaa Lys	676
ttt Phe	gat Asp	gaa Glu 220	ggt Gly	ggt Gly	gat Asp	gaa Glu	att Ile 225	gat Asp	gaa Glu	gat Asp	cta Leu	tca Ser 230	tcg Ser	tgg Trp	att Ile	724
cgc Arg	cat His 235	tcc Ser	ttg Leu	gat Asp	ctt Leu	cct Pro 240	ctt Leu	cat His	tgg Trp	agg Arg	gtc Val 245	caa Gln	gga Gly	tta Leu	gag Glu	772
gca Ala 250	aga Arg	tgg Trp	ttc Phe	tta Leu	gat Asp 255	gct Ala	tat Tyr	gcg Ala	agg Arg	agg Arg 260	ccg Pro	gac Asp	atg Met	aat Asn	cca Pro 265	820
ctt Leu	att Ile	ttc Phe	aaa Lys	ctc Leu 270	gcc Ala	aaa Lys	ctc Leu	aac Asn	ttc Phe 275	aat Asn	att Ile	gtt Val	cag Gln	gca Ala 280	aca Thr	868
tat Tyr	caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 290	agg Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	tcg Ser	tgc Cys	916
ctt Leu	gct Ala	gag Glu 300	aaa Lys	ctc Leu	cca Pro	ttt Phe	gtg Val 305	aga Arg	gat Asp	agg Arg	att Ile	gtg Val 310	gaa Glu	tgc Cys	ttc Phe	964 .
ttt Phe	tgg Trp 315	gcc Ala	atc Ile	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	a ga Ar g	1012
aaa Lys 330	atg Met	gcc Ala	gcc Ala	gtt Val	att Ile 335	att Ile	act Thr	ttc Phe	ata Ile	aca Thr 340	att Ile	atc Ile	gat Asp	gat Asp	gtt Val 345	1060
tat Tyr	gat Asp	gtg Val	Tyr	gga Gly 350	aca Thr	ata Ile	gaa Glu	Glu	cta Leu 355	gaa Glu	cta Leu	tta Leu	aca Thr	gat Asp 360	atg Met	1108
att Ile	cgc Arg	AEG	tgg Trp 365	gat Asp	aat Asn	aaa Lys	Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	Pro	tat Tyr 375	tat Tyr	atg Met	1156
caa	gtg	tgc	tat	ttg	gca	cta	tac	aac	ttc	gtt	tct	gag	cgg	gct	tac	1204

Gln	Val	Cys 380	Tyr	Leu	Ala	Leu	Tyr 385	Asn	Phe	Val	Ser	Glu 390	Arg	Ala	Tyr	
gat Asp	att Ile 395	cta Leu	aaa Lys	gat Asp	caa Gln	cat His 400	ttc Phe	aac Asn	agc Ser	atc Ile	cca Pro 405	tat Tyr	tta Leu	cag Gln	ag a Ar g	1252
tcg Ser 410	tgg Trp	gta Val	agt Ser	ttg Leu	gtt Val 415	gaa Glu	gga Gly	tat Tyr	ctt Leu	aag Lys 420	gag Glu	gca Ala	tac Tyr	tgg Trp	tac Tyr 425	1300
tac Tyr	aat Asn	ggc Gly	tat Tyr	aaa Lys 430	cca Pro	agc Ser	ttg Leu	gaa Glu	gaa Glu 435	tat Tyr	ctc Leu	aac Asn	aac Asn	gcc Ala 440	aag Lys	1348
att Ile	tca Ser	ata Ile	tcg Ser 445	gct Ala	cct Pro	aca Thr	atc Ile	ata Ile 450	tcc Ser	cag Gln	ctt Leu	tat Tyr	ttt Phe 455	aca Thr	tta Leu	1396
gca Ala	aac Asn	tcg Ser 460	att Ile	gat Asp	gaa Glu	aca Thr	gct Ala 465	atc Ile	gag Glu	agc Ser	ttg Leu	tac Tyr 470	caa Gln	tat Tyr	cat His	1444
		ctt Leu														1492
ggg Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1540
cag Gln	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	gag Glu	aga Arg	gag Glu	gcg Ala	gtg Val 520	gaa Glu	1588
cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg [°] Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1636
aca Thr	aca Thr	gcc Ala 540	agc Ser	gat Asp	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat Asp	ttg Leu	gtt Val 550	gcg Ala	gcc Ala	gca Ala	1684
gct Ala	aat Asn 555	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	Gly ggg	gat Asp	gly ggg	1732
cat His 570	Gl y	gtg Val	caa Gln	cac His	tca Ser 575	gaa Glu	ata Ile	cat His	caa Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	ctg Leu	cta Leu 585	1780
ttc Phe	cag Gln	cct Pro	tat Tyr	gtct	gaat	aa a	tcga	aaat	c ca	acct	acta	tgt	atco	ctc		1832
gata	atat	at t	ctto	gggt	t aa	cato	itt ta	att	aaaç	ittc	taat	tdaa	ag a	ıgatç	gaatcg	1892
atco	tcaa	aa a	aaaa	aaaa	ıa											1912

<210> 44

<211> 589

<212> PRT

<213> Artificial Sequence

<400> 44

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro 35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Ile Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 135 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 230 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe

290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 375 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu
405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 455 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
465 470 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu
485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 535 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr 580 585

<210> 45 <211> 1912 <212> DNA

<213	3> A1	tifi	cial	l Sec	quenc	:e										
<220 <223	3> De			on of							nce					
<221 <222 <223	<220> <221> CDS <222> (26)(1792) <223> computer-generated nucleic acid sequence encoding															
															52	
		-											tca Ser			100
													gca Ala			148
													ggg Gly 55			196
													aac Asn			244
													att Ile			292
													cag Gln			340
													ttt Phe			388
gag Glu	att Ile	aaa Lys	caa Gln 125	atc Ile	tta Leu	agt Ser	tct Ser	ata Ile 130	cac His	aat Asn	gag Glu	ccc Pro	aga Arg 135	tat Tyr	ttc Phe	436
													atc Ile			484
caa Gln	cat His 155	ggt Gly	ttt Phe	aat Asn	gtt Val	tcc Ser 160	gaa Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile	532
gag Glu 170	aag Lys	tgc Cys	agt Ser	gat Asp	ttc Phe 175	aat Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	caa Gln	gat Asp	acg Thr	aag Lys	gga Gly 185	580

atg Met	tta Leu	caa Gln	ctt Leu	tat Tyr 190	gaa (Glu)	gca Ala	tct Ser	ttc Phe	ctt Leu 195	ttg Leu	aga Arg	gaa Glu	ggt Gly	gaa Glu 200	gat Asp	628
aca Thr	ttg Leu	gag Glu	cta Leu 205	gca Ala	aga Arg	cga Arg	ttt Phe	tcc Ser 210	acs Thr	aga Arg	tct Ser	Leu	cga Arg 215	gaa Glu	aaa Lys	676
ttt Phe	gat Asp	gaa Glu 220	ggt Gly	ggt Gly	gat Asp	Glu	att Ile 225	gat Asp	gaa Glu	gat Asp	cta Leu	tca Ser 230	tcg Ser	tgg Trp	att Ile	724
cgc Arg	cat His 235	tcc Ser	ttg Leu	gat Asp	Leu	cct Pro 240	ctt Leu	cat His	tgg Tzp	agg Arg	gtc Val 245	caa Gln	gga Gly	tta Leu	gag Glu	772
gca Ala 250	aga Arg	tgg Trp	ttc Phe	tta Leu	gat Asp 255	gct Ala	tat Tyr	gcg Ala	agg Arg	agg Arg 260	ccg Pro	gac Asp	atg Met	aat Asn	cca Pro 265	820
ctt Leu	att Ile	ttc Phe	aaa Lys	ctc Leu 270	gcc Ala	aaa Lys	ctc Leu	aac Asn	tts Phe 275	aat Asn	att Ile	gtt Val	cag Gln	gca Ala 280	aca Thr	868
tat Tyr	caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 290	agg Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	tcg Ser	tgc Cys	916
ctt Leu	gct Ala	gag Glu 300	Lys	ctc Leu	cca Pro	ttt Phe	gtg Val 305	aga Arg	gat Asp	agg Arg	att Ile	gtg Val 310	gaa Glu	tgc Cys	ttc Phe	964
ttt Phe	tgg Trp 315	Ala	atc	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	aga Arg	1012
aaa Lys 330	Met	gcc	gcc Ala	gtt Val	att Ile 335	att Ile	act Thr	ttc Phe	ata Ile	aca Thr 340	att Ile	atc Ile	gat Asp	gat Asp	yal Val 345	1060
tat Tyr	gat Asp	gtg Val	tat Tyr	gga Gly 350	Thr	ata Ile	gaa Glu	gaa Glu	cta Leu 355	GIU	cta Leu	tta Leu	aca Thr	gat Asp 360	atg Met	1108
att Ile	cgc Arg	aga Arç	tgg Trp 365	Asp	aat Asn	aaa Lys	tca Ser	ata Ile 370	: Ser	caa Glr	ctt Leu	cca Pro	tat Tyi 375	: TA1	atg Met	1156
ca: Glr	gtg Val	tgc . Cy:	Ty:	ttg Lev	gca Ala	cta Leu	tac Tyr 385	ASI	tto Phe	gtt Val	tct L Ser	gaq Glu 390	, AL	g gci	t tac a Tyr	1204
gat Asj	att 7 Ile 395	Le	a aaa u Lys	a gat 3 Asp	caa Gln	cat His	Phe	aac Ası	ago n Sei	ato	e Pro	LAI	t tt: r Le	a cadu Gl	g aga n Arg	1252
tc: Se: 41	r Trį	g gt. o Va	a agi 1 Se	t tto c Lev	g gtt 1 Val 415	. Gli	a gga	a tar y Ty:	t cti	t aad u Ly: 42	s GI	g gca 1 Ala	a ta a Ty	c tg r Tr	g tac p Tyr 425	1300

tac Tyr	aat Asn	Gly	tat Tyr	aaa Lys 430	cca Pro	agc Ser	ttg Leu	gaa Glu	gaa Glu 435	tat Tyr	ctc Leu	aac Asn	aac Asn	gcc Ala 440	Lys	1348
att Ile	tca Ser	ata Ile	tcg Ser 445	gct Ala	cct Pro	aca Thr	atc Ile	ata Ile 450	tcc Ser	cag Gln	ctt Leu	tat Tyr	ttt Phe 455	aca Thr	tta L e u	1396
gca Ala	aac Asn	tcg Ser 460	att Ile	gat Asp	gaa Glu	aca Thr	gct Ala 465	atc Ile	gag Glu	agc Ser	ttg Leu	tac Tyr 470	caa Gln	tat Tyr	cat His	1444
aac Asn	ata Ile 475	ctt Leu	tac Tyr	cta Leu	tca Ser	gga Gly 480	acc Thr	ata Ile	tta Leu	agg Arg	ctt Leu 485	gct Ala	gac Asp	gat Asp	ctt Leu	1492
999 Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1540
Cag Gln	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	gag Glu	aga Arg	gag Glu	gcg Ala	gtg Val 520	gaa Glu	1588
Cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1636
aca Thr	aca Thr	gcc Ala 540	agc Ser	gat Asp	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat Asp	ttg Leu	gtt Val 55 0	gcg Ala	gcc Ala	gca Ala	1684
gct Ala	aat Asn 555	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	ggg Gly	gat Asp	ggg ggg	1732
Cat His 570	ggc Gly	gtg Val	caa Gln	cac His	tca Ser 575	gaa Glu	ata Ile	cat His	Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	ctg Leu	cta Leu 585	1780
ttc Phe	Cag Gln	cct Pro	tat Tyr	gtct	gaat	aa a	tcga	aaat	с са	acct	acta	tgt	atco	ctc		1832
						catg	ttta	att	aaag	ttc	taat	tdaa	ag a	gctg	aatcg	1892
atco	tcaa	aa a	aaaa	aaaa	a											1912

<210> 46 <211> 589 <212> PRT <213> Artificial Sequence

<400> 46

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro 35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Val Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 230 235

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

- Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380
- Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400
- Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu
 405
- Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 430
- Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445
- Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 455 460
- Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
 475 475 480
- Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 490 495
- Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510
- Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525
- Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 535 540
- Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555 560
- Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575
- Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr 580 585
- <210> 47
- <211> 1912
- <212> DNA
- <213> Artificial Sequence
- <220>
- <223> Description of Artificial Sequence:
 computer-generated sequence
- <220>
- <221> CDS
- <222> (26)..(1792)
- <223> computer-generated nucleic acid sequence encoding

(+)-sabinene synthase

<400> 47 agcaatatta caactaacaa taaaa atg tet tee att age ata aac ata get 52 Met Ser Ser Ile Ser Ile Asn Ile Ala atg cca ctg aat tcc ctc cac aac ttt gag agg aaa cct tca aaa gca Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala tgg tet acc tet tge act gca ece gca get ege ete egg gca tet tee Trp Ser Thr Ser Cys Thr Ala Pro Ala Ala Arg Leu Arg Ala Ser Ser tcc tta caa caa gaa aaa cct cac caa atc cga cgc tct ggg gat tac Ser Leu Gln Gln Glu Lys Pro His Gln Ile Arg Arg Ser Gly Asp Tyr 196 caa ccc tet ett tgg gat tte aat tae ata cag tet ate aac act ccg Gln Pro Ser Leu Trp Asp Phe Asn Tyr Ile Gln Ser Ile Asn Thr Pro 244 65 tat aag gag cag aga cac ttt aat agg caa gca gag ttg att atg caa Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ála Glu Leu Ile Met Gln 292 gtg agg atg ttg ctc aag gta aag atg gag gca att caa cag ttg gag Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu ttg att gat gac ttg caa tac ctg gga ctg tct tat ttc ttt caa gat Leu Ile Asp Asp Leu Gln Tyr Leu Gly Leu Ser Tyr Phe Phe Gln Asp 388 110 115 gag att aaa caa atc tta agt tct ata cac aat gag ccc aga tat ttc Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe 436 130 cac aat aat gat ttg tat ttc aca gct ctt gga ttc aga atc ctc aga His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg 484 caa cat ggt ttt aat gtt tcc gaa gat gta ttt gat tgt ttc aaa att Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile 532 gag aag tgc agt gat ttc aat gca aac ctt gct caa gat acg aag gga Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly 580 180 atg tta caa ctt tat gaa gca tct ttc ctt ttg aga gaa ggt gaa gat Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp 628 aca ttg gag cta gca aga cga ttt tcc acc aga tct cta cga gaa aaa Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys 676 210 ttt gat gaa ggt ggt gat gaa att gat gaa gat cta tca tcg tgg att

Phe	: Asp	Glu 220	Gly	, GJ?	/ Asp	Glu	11e 225	Asp	Glu	ı Asp	Let	3er 230		Try	lle	
C gc A rg	: cat His 235	. Jer	ttg Leu	gat Asp	ctt Leu	cct Pro 240	Leu	cat His	tgg Trp	agg Arg	gto Val 245	. Gln	gga	tta Lev	gag Glu	772
gca Ala 250	ura	tgg Trp	Phe	tta Leu	gat Asp 255	Ala	tat Tyr	gcg	agg Arg	agg Arg 260	Pro	gac Asp	atg Met	aat Asn	eca Pro 265	820
ctt Leu	att Ile	ttc Phe	aaa Lys	ctc Leu 270	ALB	aaa Lys	ctc Leu	aac Asn	ttc Phe 275	Asn	att	gtt Val	cag Gln	gca Ala 280	aca Thr	868
tat Tyr	caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 290	Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	Ser	tgc Cys	916
ctt Leu	gct Ala	gag Glu 300	aaa Lys	ctc Leu	cca Pro	ttt Phe	gtg Val 305	aga Arg	gat A sp	agg Arg	att Ile	gtg Val 310	gaa Glu	tgc Cys	ttc Phe	964
ttt Phe	tgg Trp 315	gcc Ala	atc Ile	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	aga Arg	1012
aaa Lys 330	atg Met	gcc Ala	gcc Ala	gtt Val	att Ile 335	att Ile	act Thr	ttc Phe	ata Ile	aca Thr 340	att Ile	atc Ile	gat Asp	gat Asp	gtt Val 345	1060
tat Tyr	gat Asp	gtg Val	tat Tyr	gga Gly 350	aca Thr	ata. Ile	gaa Glu	gaa Glu	cta Leu 355	gaa Glu	cta Leu	tta Leu	aca Thr	gat Asp 360	atg Met	1108
att Ile	c gc Arg	aga Arg	tgg Txp 365	gat Asp	aat Asn	aaa Lys	tca Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	cca Pro	tat Tyr 375	tat Tyr	atg Met	1156
caa Gln	gtg Val	tgc Cys 380	tat Tyr	ttg Leu	gca Ala	cta Leu	tac Tyr 385	aac Asn	ttc Phe	gtt Val	tct Ser	gag Glu 390	cgg Arg	gct Ala	tac Tyr	1204
	att Ile 395	cta Leu	aaa Lys	gat Asp	caa Gln	cat His 400	ttc Phe	aac Asn	agc Ser	atc Ile	cca Pro 405	tat Tyr	tta Leu	cag Gln	a ga A rg	1252
tcg Ser 410	tgg Trp	gta Val	agt Ser	ttg Leu	gtt Val 415	gaa Glu	gga Gly	tat Tyr	Leu	aag Lys 420	gag Glu	gca Ala	tac Tyr	tgg Trp	tac Tyr 425	1300
tac Tyr	aat Asn	Gly ggc	TAL	aaa Lys 430	cca Pro	agc Ser	ttg Leu	GLu	gaa Glu 435	tat Tyr	ctc Leu	aac Asn	aac Asn	gcc Ala 440	aag Lys	1348
att Ile	tca Ser	***	tcg Ser 445	gct Ala	cct Pro	aca :	тте .	ata Ile 450	tcc Ser	cag Gln	ctt Leu	Tyr	ttt Phe 455	aca Thr	tta Leu	1396
gca	aac	tcg	att	gat	gaa	aca (gct a	atc	gag	agc	ttg	tac	caa	tat	cat	1444

PCT/US98/20120 WO 99/15624

Ala	Asn	Ser 460	Ile	Asp	Glu	Thr	Ala 465	Ile	Glu	Ser	Leu	Tyr 470	Gln	Tyr	His	
aac Asn	ata Ile 475	ctt Leu	tac Tyr	cta Leu	tca Ser	gga Gly 480	acc Thr	ata Ile	tta Leu	agg Arg	ctt Leu 485	gct Ala	gac Asp	gat Asp	ctt Leu	1492
ggg Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1540
cag Gln	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	gag Glu	aga Arg	gag Glu	gcg Ala	gtg Val 520	gaa Glu	1588
Cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1636
aca Thr	aca Thr	gcc Ala 540	agc Ser	gat Asp	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat Asp	ttg Leu	gtt Val 550	gcg Ala	gcc Ala	gca Ala	1684
gct Ala	aat Asn 5 55	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	G] À āāā	gat Asp	gly ggg	1732
cat His 570	G ly ggc	gtg Val	Caa Gln	cac His	tca Ser 575	gaa Glu	ata Ile	cat His	caa Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	ctg Leu	cta Leu 585	1780
ttc Phe	cag Gln	cct Pro	tat Tyr	gtct	gaat	aa a	tcga	aaat	с са	acct	acta	tgt	atcc	ctc		1832
gata	atat	at t	cttg	gggt	t aa	catg	ttta	att	aaag	ttc	taat	tdaa	ag a	gctg	aatcg	18,92
atcc	tcaa	aa a	aaaa	aaaa	a											1912
<210: <211: <212: <213:	> 58: > PR'	r	cial	Seq	uenc	e										
<400: Met		Ser :	Ile :	Ser :	Ile <i>i</i>	Asn :	Ile i	Ala 1	Met 1	Pro :	Leu A	Asn :	Ser 1	Leu 1	łi s	

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro 35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Ile Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 235 230 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys
260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr
450 455 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His As: Ile Leu Tyr Leu Ser Gly
475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 495 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Glm Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 555 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr 580 585

<210> 49

<211> 1912

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (26)..(1792)

<223> computer-generated nucleic acid sequence encoding
(+)-sabinene synthase

<400> 49

agcaatatta caactaacaa taaaa atg tct tcc att agc ata aac ata gct 52
Met Ser Ser Ile Ser Ile Asn Ile Ala

atg cca ctg aat tcc ctc cac aac ttt gag agg aaa cct tca aaa gca 100
Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala
15 20 25

tg: Tr	g tet Sei	aco Thi	tct Ser	tgo Cys 30	Thr	gca Ala	Pro	gca Ala	gct Ala 35	a Arg	cto Leu	cgg Arg	gca Ala	tct Ser 40	tcc Ser	148
Se	tta Lev	Cae Gln	Caa Gln 45	GIU	aaa Lys	Pro	Cac His	Gln 50	Ile	cga Arg	cgc Arg	tct Ser	999 Gly 55	Asp	tac Tyr	196
Cas Glr	Pro	Ser 60	Leu	tgg Trp	gat Asp	ttc Phe	aat Asn 65	Tyr	ata Ile	cag Gln	tct Ser	ctc Leu 70	Asn	act Thr	ccg Pro	244
tat Tyr	Lys 75	GLU	Cag Gln	aga Arg	cac His	ttt Phe 80	aat Asn	agg Arg	caa Gln	gca Ala	gag Glu 85	gtg Val	att Ile	atg Met	caa Gln	292
90	Arg	met	Leu	Leu	Lys 95	Val	Lys	Met	Glu	gca Ala 100	Ile	Gln	Gln	Leu	Glu 105	340
neu	116	Asp	Asp	Leu 110	Gln	Tyr	Leu	Gly	Leu 115	tct Ser	Tyr	Phe	Phe	Gln 120	Asp	388
GIU	TTE	гÀ2	125	Ile	Leu	Ser	Ser	Ile 130	His	aat Asn	Glu	Pro	Arg 135	Tyr	Phe	436
urs	ASI	140	Asp	Leu	Tyr	Phe	Thr 145	Ala	Leu	gga Gly	Phe	Arg 150	Ile	Leu	Arg	484
GIII	155	GTĀ	rne	Asn	Val	Ser 160	Glu	Asp	Val	ttt Phe	Asp 165	Cys	Phe	Lys	Ile	532
170	Lys	cys	ser	Asp	Phe 175	Asn	Ala	Asn	Leu	gct Ala 180	Gln	Asp	Thr	Lys	Gly 185	580
1100	Deu	GIN	Leu	190	GLu	Ala	Ser	Phe	Leu 195	ttg Leu	Arg	Glu	Gly	Glu 200	qzA	628
		GI U	205	ALZ	Arg	Arg	Phe	210	Thr	aga Arg	Ser	Leu	Arg 215	Glu	Lys	676
	р	220	GIÀ	ст	Asp	GTA	225	Asp	Glu	gat Asp	Leu	Ser 230	Ser	Trp	Ile	724
	235	001	Deu .	wsb	red	240	red	His	Trp		Val 245	Gln	Gly	Leu	Glu	772
gca Ala 250	aga Arg	tgg Trp	ttc Phe	tta Leu	gat Asp 255	gct Ala	tat Tyr	gcg Ala	Arg	agg Arg 260	ccg Pro	gac Asp	atg Met	Asn	cca Pro 265	820

ctt Leu	att Ile	ttc Phe	aaa Lys	ctc Leu 270	gcc	aae Lys	ctc Leu	aac Asn	ttc Phe 275	aat Asn	att Ile	gtt Val	cag Gln	gca Ala 280	aca Thr	868
tat Tyr	caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 290	agg Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	tcg Ser	tgc Cys	916
ctt Leu	gct Ala	gag Glu 300	aaa Lys	ctc Leu	cca Pro	ttt Phe	gtg Val 305	aga Arg	gat Asp	agg Arg	att Ile	gtg Val 310	gaa Glu	tgc Cys	ttc Phe	964
ttt Phe	tgg Trp 315	gcc Ala	atc Ile	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	aga Arg	1012
aaa Lys 330	atg Met	gcc Ala	gcc Ala	gtt Val	att Ile 335	att Ile	act Thr	ttc Phe	ata Ile	aca Thr 340	att Ile	atc Ile	gat Asp	gat Asp	gtt Val 345	1060
tat Tyr	gat As p	gtg Val	tat Tyr	gga Gly 350	aca Thr	ata Ile	gaa Glu	gaa Glu	cta Leu 355	gaa Glu	cta Leu	tta Leu	aca Thr	gat Asp 360	atg Met	1108
att Ile	ege Arg	aga Arg	tgg Trp 36 5	gat Asp	aat Asn	aaa Lys	tca Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	cca Pro	tat Tyr 375	tat Tyr	atg Met	1156
caa Gln	gtg Val	tgc Cys 380	tat Tyr	ttg Leu	gca Ala	cta Leu	tac Tyr 385	aac Asn	ttc Phe	gtt Val	tct Ser	gag Glu 390	cgg Arg	gct Ala	tac Tyr	1204
gat Asp	att Ile 395	cta Leu	aaa Lys	gat Asp	caa Gln	cat His 400	ttc Phe	aac Asn	agc Ser	atc Ile	cca Pro 405	tat Tyr	tta Leu	cag Gln	aga Arg	1252
tcg Ser 410	tgg Trp	gta Val	agt Ser	ttg Leu	gtt Val 415	gaa Glu	gga Gly	tat Tyr	ctt Leu	aag Lys 420	gag Glu	gca Ala	tac Tyr	tgg Trp	tac Tyr 425	1300
tac Tyr	aat Asn	ggc Gly	tat Tyr	aaa Lys 430	cca Pro	agc Ser	ttg Leu	gaa Glu	gaa Glu 435	tat Tyr	ctc Leu	aac Asn	aac Asn	gcc Ala 440	aag Lys	1348
att Ile	tca Ser	ata Ile	tcg Ser 445	gct Ala	cct Pro	aca Thr	atc Ile	ata Ile 450	tcc Ser	cag Gln	ctt Leu	tat Tyr	ttt Phe 455	aca Thr	tta Leu	1396
gca Ala	aac Asn	tcg Ser 460	att Ile	gat Asp	gaa Glu	aca Thr	gct Ala 465	atc Ile	gag Glu	agc Ser	ttg Leu	tac Tyr 470	caa Gln	tat Tyr	cat His	1444
aac Asn	ata Ile 475	ctt Leu	tac Tyr	cta Leu	tca Ser	gga Gly 480	acc Thr	ata Ile	tta Leu	agg Arg	ctt Leu 485	gct Ala	gac Asp	gat Asp	ctt Leu	1492
ggg Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1540

Cag Gl n	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	GIU	aga Arg	gag Glu	gcg Ala	gtg Val 520	gaa Glu	1588
cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1636
a ca Thr	aca Thr	gcc Ala 540	agc Ser	gat Asp	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat As p	ttg Leu	gtt Val 5 50	gcg Ala	gcc Ala	gca Ala	1684
gct Ala	aat Asn 555	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	GJ A BBB	gat Asp	G J A GGG	1732
cat His 570	ggc Gly	gtg Val	caa Gln		tca Ser 575	gaa Glu	ata Ile	cat His	GID	cag Gln 580	atg Met	gga Gly	ggc Gly	Leu	cta Leu 585	1780
ttc Phe	Cag Gln	cct Pro	tat Tyr	gtct	gaat	aa a	tcga	aaat	с са	acct	acta	tgt	atcc	ctc		1832
gata	atat	at t	cttg	gggt	t aa	catg	ttta	att	aaag	ttc	taat	tdaa	ag a	gctg	aatcg	1892
atcc																1912

<210> 50

<211> 589

<212> PRT

<213> Artificial Sequence

<400> 50

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 60

Asn Tyr Ile Gin Ser Leu Asn Thr Pro Tyr Lys Glu Gin Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Val Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe

130	135	140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 230 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr

	45	60					455	i				46	0				
A1 46	a I] 5	.е (3lu	Ser	Leu	Tyz 470	Gln	Туз	H1:	s As	n Il 47	e Le 5	ц Ту	r Lei	u Se	r Gly 480	
Th	r Il	e I	Leu	Arg	Leu 485	Ala	Asp	Asp	Let	ı Gl 49	y Th.	r Se	r Glr	n His	3 Gl	u Leu 5	
G1	u Ar	g G	ly	Asp 500	Val	Pro	Lys	Ala	11e 505	Gl:	n Cy:	з Ту	. Met	Asr 510		Thr	
Ası	n Al	a S 5	er 15	Glu	Arg	Glu	Ala	Val 520	Glu	Hi:	s Val	L Lys	Phe 525	Lev	Ile	Arg	
Gl	1 Al 53	а Т 0	rp	Lys	Glu	Met	Asn 535	Thr	Val	Th	r Thi	Ala 540	Ser	Asp	Cys	Pro	
•••						220					555	•				Ala 560	
					363					570)			His	Ser 575	Glų	
Ile	Hi:	5 G.	ln	Gln 580	Met	Gly	Gly	Leu	Leu 585	Ph∈	Gln	Pro	Tyr				
<21 <21 <22 <22 <22 <22 <22 <22	0> 3> D 0> 1> C 2> (912 ONA LITTI Cesc Comp DS 26)	lfi erip oute	ptio er-g	2)	Art ated	ific nuc	leic	: aci	id s	eque						
	,	17-	ute	er-g	ener	ated ynth	nuc ase	leic	aci	.d s	equei	nce e	encod	ling			
	0> 5 aata		Ca	act	acaa	a ta	aaa ,	atg Met 1	tct Ser	tcc Ser	att Ile	agc Ser 5	ata Ile	aac Asn	ata Ile	gct Ala	52
atg Met 10	cca Pro	ct Le	ga uA	at t	er i	teu l 15	cac a	aac Asn	ttt Phe	gag Glu	agg Arg 20	aaa Lys	cct Pro	tca Se <i>r</i>	aaa Lys	gca Ala 25	100
•				C	30	ine 2	ma i	PIO 1	Ala.	35	cgc Arg	Leu	Arg .	Ala	Ser 40	Ser	148
tcc Ser	tta Leu	Ca: Gl:		aa g ln G 45	aa a lu I	aa d ys I	ect o	ac (lis (Sln 50	atc Ile	cga Arg	cgc Arg	tct (ggg Gly 2 55	gat Asp	tac Tyr	196
caa	ccc	tc	t c	tt t	gg g	at t	tc a	ati	ac a	ata	cag	tct	ctc a	aac a	act	ccg	244

G1:	n Pr	O Se	er Le 50	u Tr	p As	P Phe	AST 65	ту. 5	r Il	e Gl	n Se	r Lei 70	u As: 0	n Th	r Pro	
ta: Ty:	t aa r Ly 7		ig ca .u Gl	g ag n Ar	a cad g His	c ttt s Phe 80	AST	age Are	g caa	a gc n Al	a gad a Gli 8	u Le	g ati	t at e Me	g caa t Gln	292
gte Val 90		g at g Me	g tt t Le	g ct u Le	c aaq u Lys 95	s Agt	aag Lys	ato Met	g gaq E Glu	g gc u Ala 100	a Let	caa 1 Glr	a cad	j tte	g gag 1 Glu 105	340
tt: Le:	g att	t ga e As	t ga p As	c tto P Lei 110	d GIN	tac Tyr	ctg Leu	gga	Lev 115	ı Sei	t tat	tto Phe	ttt Phe	Caa Gl: 120	a gat a Asp	388
gag Glu	att Ile	aa Ly	a cad s Gli 12:		tta Leu	agt Ser	tct Ser	ata Ile 130	His	aat Asr	gag 1 Glu	Pro	aga Arg 135	Ty	ttc Phe	436
Cac His	aat Asn	: aat Asi 140	إشامنا	t ttg	tat Tyr	ttc Phe	aca Thr 145	gct Ala	ctt Leu	gga	ttc Phe	aga Arg 150	Ile	Leu	aga Arg	484
caa Gln	cat His 155		ttt y Phe	aat Asn	gtt Val	tcc Ser 160	gaa Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile	532
gag Glu 170	aag Lys	tgc Cys	egt Ser	gat Asp	Phe 175	aat Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	Gln	gat Asp	acg Thr	aag Lys	gga Gly 185	580
atg Met	tta Leu	Caa Gln	ctt Leu	tat Tyr 190	GIU	gca Ala	tct Ser	ttc Phe	ctt Leu 195	ttg Leu	aga Arg	gaa Glu	ggt Gly	gaa Glu 200	gat Asp	628
		014	Leu 205	ALG	Arg	Arg	rne	210	Thr	Arg	Ser	Leu	Arg 215	Glu	Lys	676
		220		GIY	Asp	GIU	225	qzA	Glu	Asp	Leu	Ser 230	Ser	Trp	Ile	724
-	235		ttg Leu	nap	beu	240	ren i	Hls	Trp	Arg	Val 245	Gln	Gly	Leu	Glu	7 72
250			ttc Phe	Leu	255	ALA .	ryr /	ALA .	Arg	260	Pro	Asp	Met	Asn	Pro 265	820
			aaa Lys	270	Λια	шуз ,	Leu /	Asn	275	Asn	Ile	Val	Gln .	Ala 280	Thr	868
			gaa Glu 285		шуз ,	nap .	2	290	Arg	Trp	Trp .	Asn :	Ser 295	Ser	Cys	916
ctt	gct	gag	aaa	ctc	cca	ttt	gtg a	aga (gat	agg	att	gtg (gaa	tgc	ttc	964

Le	u Al	.a G:	lu L; 00	ys Le	u Pr	o Ph	e Va. 30	l Ar	g Ası	p Ar	g Ile	• Val	l Gl	и су	s Phe	
tt Ph	t tg e Tr 31		c at la Il	c go le Al	g gc a Al	t tt a Pho 320	e etr	g cci	t cad	Caa Glr	tat Tyr 325	: Ser	tai Tyi	c Gl	g aga n Arg	101/2
12: 13:33:		g go t Al	c go a Al	c gt .a Va	t at 1 11 33	e TTG	t act Thr	tto Phe	ata : Ile	aca Thr	: Ile	ato Ile	gat Asp	gat Asp	gtt Val 345	1060
- 3		, vu	y	35	y In.	r TTE	e GIU	GIU	355	Glu	Leu	Leu	Thr	360		1108
		,	36	5 AS	P ASI	t aaa 1 Lys	ser	370	Ser	Gln	Leu	Pro	Tyr 375	Tyr	Met	1156
		38	0	r Det	1 MIS	cta Leu	385	Asn	Phe	Val	Ser	Glu 390	Arg	Ala	Tyr	1204
	395	5	L LJy.	, vot	GIN	cat His 400	Pne	Asn	Ser	Ile	Pro 405	Tyr	Leu	Gln	Arg	1252
410		, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		Leu	415		GTÅ	Tyr	Leu	Lys 420	Glu	Ala	Tyr	Trp	Tyr 425	1300
			y -	430	PIO	agc Ser	Leu	GIU	435	Tyr	Leu	Asn	Asn	Ala 440	Lys	1348
			445	Ala	FIO	aca Thr	TIE	450	Ser	Gln	Leu	Tyr	Phe 455	Thr	Leu	1396
		460	- 110	- veb	GIU	aca Thr	465	IIe	Glu	Ser	Leu	Tyr 470	Gln	Tyr	His	1444
	475			neu	SEL	gga Gly 480	The	TIE	Leu	Arg	Leu . 485	Ala	Asp .	qeA	Leu	1492
49 0				.11.2	495	ctg Leu	GIU.	Arg	GIÀ .	Asp 500	Val 1	Pro 1	Lys .	Ala	Ile 505	1540
	•	•		510	тшр	aca Thr	voli 1	n_a	515	GIU .	Arg (slu A	Ma Y	Val 520	Glu	1588
			525		116	agg Arg	GIU ,	530	rrp .	Lys (Glu N	1et A	lsn : 35	Chr '	Val	1636
aca	aca	gcc	agc	gat	tgt	ccg	ttt a	acg (gat q	gat 1	ttg ç	jtt g	ica d	gcc (gca	1684

Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala 540 545 550

gct aat ctt gca agg gcg gct cag ttt ata tat ctc gac ggg gat ggg 1732 Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly 555 560 565

cat ggc gtg caa cac tca gaa ata cat caa cag atg gga ggc ctg cta
His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu
570
580
585

ttc cag cct tat gtctgaataa atcgaaaatc caacctacta tgtatccctc 1832
Phe Gln Pro Tyr

gataatatat tottggggtt aacatgttta attaaagtto taattdaaag agotgaatog 1892 atootcaaaa aaaaaaaaa 1912

<210> 52 <211> 589

<212> PRT

<213> Artificial Sequence

<400> 52

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala
20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe
65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Leu Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 225

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 255 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
475 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu
485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 520 Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 570 Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr <210> 53 <211> 1912 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <**22**2> (26)..(1792) <223> computer-generated nucleic acid sequence encoding (+)-sabinene synthase agcaatatta caactaacaa taaaa atg tct tcc att agc ata aac ata gct Met Ser Ser Ile Ser Ile Asn Ile Ala 52 atg cca ctg aat tcc ctc cac aac ttt gag agg aaa cct tca aaa gca 100 Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala 10 tgg tot acc tot tgc act gca coo gca gct cgc oto cgg gca tot tec Trp Ser Thr Ser Cys Thr Ala Pro Ala Ala Arg Leu Arg Ala Ser Ser tee tta caa caa gaa aaa eet cae caa ate ega ege tet ggg gat tae 196 Ser Leu Gln Gln Glu Lys Pro His Gln Ile Arg Arg Ser Gly Asp Tyr caa ccc tot ott tgg gat tto aat tac ata cag tot otc aac act ccg 244 Gin Pro Ser Leu Trp Asp Phe Asn Tyr Ile Gin Ser Leu Asn Thr Pro 65 tat aag gag cag aga cac ttt aat agg caa gca gag ttg att atg caa Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ala Glu Leu Ile Met Gln gtg agg atg ttg ctc aag gta aag atg gag gca att caa cag gtg gag Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Val Glu 95 100

tt: Le:	y att	gat Asp	gac Asp	ttg Leu 110	GIN	tac	ctg Leu	gga	Leu 115	Ser	tat Tyr	ttc Phe	ttt Phe	Gln 120	Asp	388
<u>e</u> fn	att Ile	Lys	Gln 125	Ile	tta Leu	agt Ser	tct Ser	ata Ile 130	cac His	aat Asn	gag Glu	ccc Pro	aga Arg 135	tat Tyr	ttc Phe	436
Cac His	aat Asn	aat Asn 140	qeA	ttg Leu	tat Tyr	ttc Phe	aca Thr 145	gct Ala	ctt Leu	gga Gly	ttc Phe	aga Arg 150	atc Ile	ctc Leu	a ga Ar g	484
caa Gln	cat His 155	GTÅ	ttt Phe	aat Asn	gtt Val	tcc Ser 160	gaa Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile	532
gag Glu 170	aag Lys	tgc Cys	agt Ser	gat Asp	ttc Phe 175	aat Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	caa Gln	gat Asp	acg Thr	aag Lys	gga Gly 185	580
atg Met	tta Leu	caa Gln	ctt Leu	tat Tyr 190	gaa Glu	gca Ala	tct Ser	ttc Phe	ctt Leu 195	ttg Leu	a ga Arg	gaa Glu	ggt Gly	gaa Glu 200	gat Asp	628
Inc	ren	Glu	205	Ala	Arg	Arg	Phe	Ser 210	Thr	Arg	Ser	Leu	Arg 215	gaa Glu	Lys	676
rne	Asp	220	СтÀ	GIÀ	Asp	Glu	11e 225	Ąsp	Glu	Asp	Leu	Ser 230	Ser	tgg Trp	Ile	724
ALG	235	Ser	Leu	Asp	Leu	Pro 240	Leu	His	Trp	Arg	Val 245	Gl'n	Gly	tta Leu	Glu	77 2
250	Arg	тгр	rne	Leu	Asp 255	Ala	Tyr	Ala	Arg	Arg 260	Pro	Asp	Met	aat Asn	Pro 265	820
neu	TIE	rne	Lys	270	Ala	Lys	Leu	Asn	Phe 275	Asn	Ile	Val	Gln	gca Ala 280	Thr	868
-3-	4III	GIU	285	Leu	Lys	Asp	Ile	Ser 290	Arg	Trp	Trp	Asn	Ser 295	tcg Ser	Cys	916
	AL a	300	гуз	reu	Pro	Phe	Val 305	Arg	Asp	Arg	Ile	Val 310	Glu	tgc Cys	Phe	964
	315	AL C	TIE .	ALA .	Ala	Phe 320	Glu	Pro	His	Gln	Tyr 325	Ser '	Tyr	cag Gln	Arg.	1012
aaa Lys 330	atg Met	gcc Ala	gcc Ala	val	att Ile 335	att Ile	act Thr	ttc Phe	Ile	aca Thr 340	att Ile	atc Ile <i>i</i>	gat Asp	gat Asp	gtt Val 345	1060

tat Ty:	gat Asp	gto Val	tat Tyr	gga Gly 350	Thr	ata Ile	gaa Glu	gaa Glu	cta Leu 355	. Glu	cta Leu	tta Leu	aca Thi	gat Asp 360	atg Met	1108
att Ile	e Arg	aga Arg	tgg Trp 365	ASP	aat Asn	a aa Lys	tca Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	cca Pro	tat Tyr 375	Tyr	atg Met	1156
Gl n	gtg Val	Cys 380	Tyt	ttg Leu	gca Ala	cta Leu	tac Tyr 385	aac Asn	ttc Phe	gtt Val	tct Ser	gag Glu 39 0	Arg	gct Ala	tac Tyr	1204
vah	395	beu	гÀз	Asp	GIN	H15 400	Phe	Asn	Ser	Ile	Pro 405	Туг	Leu	Gln		1252
410	IIP	Val	ser	rea	Val 415	GLu	СΙΆ	tat Tyr	Leu	Lys 420	Glu	Ala	Tyr	Trp	Tyr 425	1300
• } -	ASII	GIY	Tyr	130	Pro	ser	Leu	gaa Glu	Glu 435	Tyr	Leu	Asn	Asn	Ala 440	Lys	1348
TIE	ser	iie	445	Ala	Pro	Thr	Ile	ata Ile 450	Ser	Gln	Leu	Tyr	Phe 455	Thr	Leu	1396
Λια	ASII	460	TTG	Asp	GLu	Thr	Ala 465	atc Ile	Glu	Ser	Leu	Tyr 470	Gln	Tyr	His	1444
wo II	475	Leu	TYE	rea	Ser	G1y 480	Thr	ata Ile	Leu	Arg	Leu 485	Ala	Asp	Asp	Leu	1492
490	Int	ser	GIN	Hls	495	Leu	Glu	aga Arg	Gly	Asp 500	Val	Pro	Lys	Ala	11e 505	1540
	O,S	TYL	riec	510	Asp	Tnr .	Asn		Ser 515	Glu	Arg	Glu	Ala	Val 520	Glu	1588
	·	шуз	525	Leu	IIG .	Arg	GIU .	gcg Ala 530	Тгр	Lys	Glu	Met	Asn 535	Thr	Val	1636
		540	Ser	Asp	Cys	PEO	545	acg Thr	Азр	Asp	Leu	Val 550	Ala	Ala	Ala	1684
	555			ALG	A14 /	560	3TU	ttt . Phe	TIE	Tyr	565	Asp	Gly	Asp	Gly.	1732
Cat His 570	GJ Y Gg C	gtg Val	caa Gln	ura	tca (Ser (57 5	gaa (Glu)	ata Ile i	cat His	Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	Leu	cta Leu 585	1780

ttc cag cct tat gtctgaataa atcgaaaatc caacctacta tgtatccctc 1832 Phe Gln Pro Tyr

gataatatat tottggggtt aacatgttta attaaagtto taattdaaag agotgaatog 1892

atcctcaaaa aaaaaaaaaa

1912

<210> 54

<211> 589

<212> PRT

<213> Artificial Sequence

<400> 54

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Val Glu Leu Ile Asp Asp Leu Gln Tyr
100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 155 1560

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 235 230 235

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp
275
280
285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 375 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 450

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
465 470 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu
485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 535 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555 560

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu
565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr
580 585

<210> 55

<211> 1912

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (26)..(1792)

<223> (+)-sabinene synthase encoded by computer
generated nucleic acid sequence

<400> 55

agcaatatta caactaacaa taaaa atg tct tcc att agc ata aac ata gct 52

Met Ser Ser Ile Ser Ile Asn Ile Ala
1

atg cca ctg aat tcc ctc cac aac ttt gag agg aaa cct tca aaa gca 100
Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala
10 20

tgg tct acc tct tgc act gca ccc gca gct cgc ctc cgg gca tct tcc 148
Trp Ser Thr Ser Cys Thr Ala Pro Ala Ala Arg Leu Arg Ala Ser Ser
30 35

tcc tta caa caa gaa aaa cct cac caa atc cga cgc tct ggg gat tac 196 Ser Leu Gln Glu Lys Pro His Gln Ile Arg Arg Ser Gly Asp Tyr 45

Caa ecc tet ett tgg gat tte aat tae ata eag tet ete aac act eeg 244
Gln Pro Ser Leu Trp Asp Phe Asn Tyr Ile Gln Ser Leu Asn Thr Pro
60 65 70

tat aag gag cag aga cac ttt aat agg caa gca gag ttg att atg caa 292
Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ala Glu Leu Ile Met Gln
75 80 85

gtg agg atg ttg ctc aag gta aag atg gag gca att caa cag ttg gag 340 Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu 90 95 100 105

gtg att gat gac ttg caa tac ctg gga ctg tct tat ttc ttt caa gat 388
Val Ile Asp Asp Leu Gln Tyr Leu Gly Leu Ser Tyr Phe Phe Gln Asp
110 115

gag att aaa caa atc tta agt tct ata cac aat gag ccc aga tat ttc 436 Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe 125 130 135

cac aat aat gat ttg tat ttc aca gct ctt gga ttc aga atc ctc aga 484

			Asp					•					120				
	155		ttt Phe			160		امد	y va	T P1	1e ,1	.65	Cys	Phe	Lys	Ile	
170			agt Ser	•	175				, ne	18	10	ıın A	usp :	Thr	Lys	Gly 185	580
			ctt :	190					19	5 5	u A	rg G	TH (31 y	Glu 200	Asp	628
			cta d Leu <i>l</i> 205			- 3		210	1111	· AL	g se	er L	eu A 2	15	Glu	Lys	676
Phe i	;	220		_	•	2	225	. wp	GIU	. Asi	э ге	2:	er S 30	er :	l, rp	Ile	724
Arg F	235			_	2	10			тър	ΑLC	24	5	n G	ly I	eu	Glu	772
gca a Ala A 250				25	5	_	,	<u></u> a	ALG	260	PE	o As	P Me	et A	sn i	Pro 265	820
ctt a Leu I			27	70	-2			~ II	275	ASI	110	e va	I G1	.n A 2	la 1 80	hr	868
tat c		2	85	-			2	90	nry	ттр	Trį	Ası	n Se 29	r 5 5	er C	ys	916
Leu A	30	00				30	5	-9 /	чp	Arg	TIE	310) r et	u c	/S P	he	964
ttt tg Phe Tr 31	.5				320)				GIN	325	ser	Ty	r Gl	n A	rg	1012
aaa at Lys Me 330				335				1	16	340	тте	IIe	Asp) As	p Va 34	11 15	1060
tat ga Tyr As			350)				3	55	JLU	Leu	Leu	Inr	36	t at p Me	g t	1108
att cg	s ag	a tg g Tr 36	g gat P Asp 5	aat Asn	aaa Lys	tc: Se:	a at r Il 37		gc d	aa 31n 3	ctt Leu	cca Pro	tat Tyr 375	ta:	tat r Me	g t	1156
caa gto	; tg	c ta	ttg	gca	cta	tac	a a	c ti	tc g	tt (ct	gag		gct	: ta	c	1204

G1	ln V	al c	ys 80	Tyr	Lei	ı Al	a Le	u T 3	yr . 85	Asn	Ph	e V	al s	er	Glu 390	Ar	g Ai	la :	Tyr		
	3	95			•	-	40	-		w.,	36	٠	1e P	05	ryr	Le	n C]	ln A	lrg	12	5⊉
41	0					41	5	a gç u Gl		· y ~	ne.	42	75 G. 20	ıu A	чта	ТУ	Tr	P 1	yr 25	13	00
				_	430			c tt r Le			435	1 1y	L Le	eu A	ısn	Asn	44	a L O	ys	134	48
			4	45				a at	4	50	3e I	GI	n Le	eu T	yr	Phe 455	Th:	r L	eu	139	6
		46	0		•			465	5		GIU	se.	r re	4	yr 70	Gln	Туі	r H:	ĹS	144	4
	475	i					480			1	ueu	Arg	48.	u A. 5	la /	\sp	Asp	Le	eu.	149	2
490						495		gag Glu		9 (ътÀ	500) va.	LPI	co I	.ys	Ala	11 50	.e	154	0
				5	10	•		aat Asn		5	15	GIU	Arg	, GT	u A	ца	Val 520	G1	u	1588	3
			52	5				gag Glu	53	ō	-1	Dy S	GIU	. Me	5 A	35	Thr	Va.	1	1636	5
		540			•			ttt Phe 545			sp.	Asp	reu	55	0 1 A	la Z	Ala	Ala	3	1684	•
•	555				•		560	cag Gln		- 4.		1 Y L	565	Ası	pG.	Ly A	Asp	Gl	1	1732	
is (el y ggc	gtg Val	Caa Gln	ca Hi		ca er 75	gaa Glu	ata Ile	cat His	Ca G1		cag Gln 580	atg Met	gga Gly	a gg	y I	.eu	cta Leu 585	ı	1780	
tc d	ag Sln	cct Pro	tat Tyr	gt	ctg	aat	aa a	tcga	aaa	tc	caa	acct	acta	a tg	gtat	ccc	tc			1832	
ataa	tat	at t	ctt	9 99	gtt	aad	atg	ttta	at	taa	agt	tc	taat	tda	aag	ag	ctg	aat	cg	1892	
teet	caa	aa a	aaa	aaa.	aaa															1912	

<210> 56

<211> 589

<212> PRT

<213> Artificial Sequence

<400> 56

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

1 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala
20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Val Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 230 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe

300

290 295

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu
405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 455 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
465 470 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555 560

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr
580 585

<210> 57 <211> 1912 <212> DNA

<2	213>	Art:	Lfic	ial S	ieque	ence										
	20> 23>	Desc	ript	ion gen	of A	urtificed n	icia ucle	l Se	equen	sequ	lence	.				
<2 <2	23>	CDS (26) comp	(1	.792) :-gen	erat	ed n	ucle						odin	g		
	00> caat		C2.2	ctaa		-										
							Me	l l	r se	: I1	e Se	r Ile 5	e Ası	ı Il	a gct e Ala	52
10)				15	5	, Asi	ı PM	e GI	20	g Ly:	s Pro	Se:	Ly	a gca s Ala 25	100
-			- 55,	30)	. A.	PIC	AL	35	Arg	J Let	ı Arg	Ala	Se:		148
			45	5	. .	210	urs	50) r TTG	Arg	J Arg	, Ser	Gly 55	Asp	tac Tyr	196
		60			rap	rne	65	ryr	ile	Gln	Ser	70	Asn	Thr	ccg Pro	244
tat Tyr	aag Lys 75	gag Glu	Gln	aga Arg	cac His	ttt Phe 80	aat Asn	agg Arg	caa Gln	gca Ala	gag Glu 85	ttg Leu	att Ile	atg Met	caa Gln	292
gtg Val 90	agg Arg	atg Met	ttg Leu	ctc Leu	aag Lys 95	gta Val	aag Lys	atg Met	gag Glu	gca Ala 100	att Ile	caa Gln	cag Gln	ttg Leu	gag Glu 105	340
ttg Leu	att Ile	gat Asp	gac Asp	gtg Val 110	caa Gln	tac Tyr	ctg Leu	gga Gly	ctg Leu 115	tct Ser	tat Tyr	ttc Phe	ttt Phe	caa Gln 120	gat Asp	388
gag Glu	att Ile	a aa Lys	caa Gln 125	atc Ile	tta Leu	agt Ser	tct Ser	ata Ile 130	cac His	aat Asn	gag Glu	ccc Pro	aga Arg 135	tat Tyr	ttc Phe	436
cac His	aat Asn	aat Asn 140	gat As p	ttg Leu	tat Tyr	ttc Phe	aca Thr 145	gct Ala	ctt Leu	gga Gly	ttc Phe	aga Arg 150	atc Ile	ctc Leu	aga Arg	484
caa Gln	cat His 155	ggt Gly	ttt Phe	aat Asn	gtt Val	tcc Ser 160	gaa Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile -	532
gag Glu 170	aag Lys	tgc Cys	agt Ser	gat Asp	ttc Phe 175	aat Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	caa Gln	gat Asp	acg Thr	aag Lys	gga Gly 185	580

				190		, A.	. Jei	. Pne	19	u Leu 5	Arg	gaa Glu	Gly	Glu 200	Asp	628
			205	,	,,,,	ALY	File	210)	e Arg	Ser	cta Leu	Arg 215	Glu	Lys	676
•	•	220	3	,	-1.00	GIU	225	Asp	GIT	ı Asp	Leu	tca Ser 230	Ser	Trp	Ile	724
	235			·p	Deu	240	neu	HIS	Trp	Arg	Val 245	caa Gln	Gly	Leu	Glu	772
250	_	•			255	ALG	TYL	Λια	Arg	260	Pro	gac Asp i	Met	Asn	Pro 265	820
			-,-	270	744	Буз	Leu	Asn	275	Asn	Ile	gtt (Val (Gln .	Ala 280	Thr	868
			285		- ,5	nsp	116	290	Arg	Trp	Trp .	2	er :	Ser	Cys	916
Ctt (300	•			:	305	ALG	Asp	Arg	ile :	Val G 310	ilu (Cys 1	Phe	964
	115					320	JIU	FEO	nis	GIN	Tyr :	ser T	yr G	iln)	yi rd	1012
aaa a Lys M 330		-		3	335	LIE ,	ine i	rne	TTE	340	lle]	lle A	sp A	sp V 3	/al 845	1060
tat g Tyr A	_			350		ile c	Lu (31 U .	355	GIU I	Seu I	eu T	hr A 3	sp M 60	let	1108
att c Ile A	_	3	65	·	~	ys s	3	70	ser (GIN I	.eu P	ro Ty	72 T	yr M	let	1156
Caa gi Gln Va	3	80	-			3	85	sn E	ne v	val S	er G	lu Ar 90	g A	la T	уr	1204
gat at Asp II	95		•	- P •	4	00	ne v	an a	er 1	11e P	05	yr Le	u G	ln A	rg -	1252
ser Tr 410	gg gt	ta a al S	gt t er L		tt g al G 15	aa ge lu G	ga t ly T	at c yr L	eu I	ag g ys G 20	ag g lu A	ca ta la Ty	c to	p T	ac yr 25	1300

-•		,	-,-	430	PEC) Ser	Leu	GI	435	Tyr	Leu	Asn	Asn	Ala 440		1348
			445	Λια	PEC	Thr	11e	450	Ser	Gln	Leu	Tyr	Phe 455	Thr	tta Leu	1396
		460	176	Asp	GIU	rnr	465	Ile	Glu	Ser	Leu	Tyr 470	Gln	Tyr		1444
	475	Deu	. 7.	Deu	ser	gga Gly 480	THE	IIe	Leu	Arg	Leu 485	Ala	Asp	Asp	Leu	1492
490			0111	NIS	495	ctg Leu	GIU	Arg	GIÀ	Asp 500	Val	Pro	Lys	Ala	11e 505	1540
	-,,	-7-	.nec	510	ASP	aca Thr	Asn	ALA	515	Glu	Arg	Glu	Ala	Val 520	Glu	1588
		2,5	525	Deu	116	agg Arg	GIU	530	Trp	Lys	Glu	Met	Asn 535	Thr	Val	1636
		540	Der	Asp	Cys		545	Thr	Asp	Asp	Leu	Val 550	Ala	Ala	Ala.	1684
	555		ALC.	Arg	ALA	gct Ala 560	GIN	Phe	Ile	Tyr	Leu 565	Asp	Gly .	Asp	Gly	1732
570	1	V	GIII .	urs	575	gaa Glu	TTG	Hls	Gln	Gln 1 580	Met	Gly (Gly :	Leu	cta Leu 585	1780
ttc Phe	cag Gln	cct : Pro :	tat (Tyr	gtct	gaat	aa a	tcga	aaat	c ca	acct	acta	tgta	atce	ctc	•	1832
						catg	ttta	att	aaag	ttc :	taati	tdaaa	ag ag	gctg	aatcg	1892
atcc	tcaa.	aa aa	aaaa	aaa	a											1912

<210> 58

<211> 589 <212> PRT

<213> Artificial Sequence

<400> 58

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro 35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Val Gln Tyr
100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 135 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 230 235 240

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 345 350

Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr <210> 59 <211> 1912 <212> DNA <213> Artificial Sequence

<223> Description of Artificial Sequence: computer-generated nucleic acid sequence

<220>

<220>

<221> CDS

<222> (26) .. (1792)

<223> computer-generated nucleic acid sequence encoding

(+)-sabinene synthase

<400 agca	0> 59 satat) :ta c	aact	aaca	a ta	aaa	atg Met 1	tct Ser	tcc Ser	att Ile	agc Ser 5	ata Ile	aac Asn	ata Ile	gct Ala	52
atg Met 10	cca Pro	ctg Leu	aat Asn	tcc Ser	ctc Leu 15	cac His	aac Asn	ttt Phe	gag Glu	agg Arg 20	aaa Lys	cct Pro	tca Ser	aaa Lys	gca Ala 25	100
tgg Trp	tct Ser	acc Thr	tct Ser	tgc Cys 30	act Thr	gca Ala	ccc Pro	gca Ala	gct Ala 35	cgc Arg	ctc Leu	cgg Arg	gca Ala	tct Ser 40	tcc Ser	148
tcc Ser	tta Leu	caa Gln	caa Gln 45	gaa Glu	aaa Lys	cct Pro	cac His	caa Gln 50	atc Ile	cga Arg	cgc Arg	tct Ser	ggg Gly 55	gat Asp	tac Tyr	196
caa Gln	ccc Pro	tct Ser 60	ctt Leu	tgg Trp	gat Asp	ttc Phe	aat Asn 65	tac Tyr	ata Ile	cag Gln	tct Ser	ctc Leu 70	aac Asn	act Thr	ecg Pro	244
tat Tyr	aag Lys 75	gag Glu	cag Gln	aga Arg	cac His	ttt Phe 80	aat Asn	agg Arg	caa Gln	gca Ala	gag Glu 85	ttg Leu	att Ile	atg Met	ca a Gl n	292
gtg Val 90	agg Arg	atg Met	ttg Leu	ctc Leu	aag Lys 95	gta Val	aag Lys	atg Met	gag Glu	gca Ala 100	att Ile	caa Gln	cag Gln	ttg Leu	gag Glu 105	340
ttg Leu	att Ile	gat Asp	gac Asp	ttg Leu 110	caa Gln	tac Tyr	gtg Val	gga Gly	ctg Leu 115	tct Ser	tat Tyr	ttc Phe	ttt Phe	caa Gln 120	gat As p	388
gag Glu	att Ile	aaa Lys	caa Gln 125	atc Ile	tta Leu	agt Ser	tct Ser	ata Ile 130	cac His	aat Asn	gag Glu	Pro	aga Arg 135	tat Tyr	ttc Phe	436
cac His	aat Asn	aat Asn 140	gat Asp	ttg Leu	tat Tyr	ttc Phe	aca Thr 145	gct Ala	ctt Leu	gga Gly	ttc Phe	aga Arg 150	Ile	ctc Leu	aga Arg	484
caa Gln	cat His 155	Gly	ttt Phe	aat Asn	gtt Val	tcc Ser 160	Glu	gat Asp	gta Val	ttt Phe	gat Asp 165	tgt Cys	ttc Phe	aaa Lys	att Ile	532
gag Glu 170	Lys	tgc Cys	agt Ser	gat Asp	ttc Phe 175	aat Asn	gca Ala	aac Asn	ctt Leu	gct Ala 180	Gln	gat Asp	acg Thr	Lys	gga Gly 185	580
atg Met	tta Leu	caa Gln	ctt Leu	tat Tyr 190	Glu	gca Ala	tct Ser	ttc Phe	ctt Leu 195	Leu	aga Arg	gaa Glu	ggt Gly	gaa Glu 200	Asp	628
aca Thr	ttg Leu	gag Glu	cta Leu 205	Ala	aga Arg	cga Arg	ttt Phe	tcc Ser 210	Thr	aga Arg	tct Ser	cta Leu	cga Arg 215	, Glu	aaa Lys	676
ttt	gat:	gaa	ggt	ggt	gat	gaa	att	gat	gaa	gat	cta	tca	tcg	, tgg	att	724

j	Phe	Asp	Glu 220	Gly	Gly	Ąsp	Glu	Ile 225	Ąsp	Glu	Asp	Leu	Ser 230	Ser	Trp	Ile	
j	ege Arg	cat His 235	tcc Ser	ttg Leu	gat Asp	ctt Leu	cct Pro 240	ctt Leu	cat His	tgg Trp	agg Arg	gtc Val 245	caa Gln	gga Gly	tta Leu	gag Glu	772 <i>‡</i>
2	gca Ala 250	aga Arg	tgg Trp	ttc Phe	tta Leu	gat Asp 255	gct Ala	tat Tyr	gcg Ala	agg Arg	agg Arg 260	ccg Pro	gac Asp	atg Met	aat Asn	Pro 265	820
:	ctt Leu	att Ile	ttc Phe	aaa Lys	ctc Leu 270	gcc Ala	aaa Lys	ctc Leu	aac Asn	ttc Phe 275	aat Asn	att Ile	gtt Val	cag Gln	gca Ala 280	aca Thr	868
•	tat Tyr	caa Gln	gaa Glu	gaa Glu 285	ctg Leu	aaa Lys	gat Asp	atc Ile	tca Ser 29 0	agg Arg	tgg Trp	tgg Trp	aat Asn	agt Ser 295	tcg Ser	tgc Cys	916
	ctt Leu	gct Ala	gag Glu 300	aaa Lys	ctc Leu	cca Pro	ttt Phe	gtg Val 305	aga Arg	gat Asp	agg Arg	att Ile	gtg Val 310	gaa Glu	tgc Cys	ttc Phe	964
	ttt Phe	tgg Trp 315	gcc Ala	atc Ile	gcg Ala	gct Ala	ttt Phe 320	gag Glu	cct Pro	cac His	caa Gln	tat Tyr 325	agt Ser	tat Tyr	cag Gln	aga Arg	1012
	aaa Lys 330	atg Met	gcc Ala	gcc Ala	gtt Val	att Ile 335	att Ile	act Thr	ttc Phe	ata Ile	aca Thr 340	att Ile	atc Ile	gat Asp	gat Asp	gtt Val 345	1060
	tat Tyr	gat Asp	gtg Val	tat Tyr	gga Gly 350	aca Thr	ata Ile	gaa Glu	gaa Glu	cta Leu 355	gaa Glu	cta Leu	tta Leu	aca Thr	gat Asp 360	atg Met	1108
	att Ile	ege Arg	aga Arg	tgg Trp 365	gat Asp	aat Asn	aaa Lys	tca Ser	ata Ile 370	agc Ser	caa Gln	ctt Leu	CCA Pro	tat Tyr 375	tat Tyr	atg Met	1156
	caa Gln	gtg Val	tgc Cys 380	tat Tyr	ttg Leu	gca Ala	cta Leu	tac Tyr 385	aac Asn	ttc Phe	gtt Val	tct Ser	gag Glu 390	cgg Arg	gct Ala	tac Tyr	1204
	gat Asp	att Ile 395	cta Leu	aaa Lys	gat Asp	caa Gln	cat His 400	Phe	aac Asn	agc Ser	atc Ile	cca Pro 405	tat Tyr	tta Leu	cag Gln	aga Arg	1252
	tcg Ser 410	Trp	gta Val	agt Ser	ttg Leu	gtt Val 415	Glu	gga Gly	tat T <u>y</u> r	ctt Leu	aag Lys 420	Glu	gca Ala	tac Tyr	tgg Trp	tac Tyr 425	1300
	tac Tyr	aat Asn	G]A ggc	tat Tyr	aaa Lys 430	Pro	agc Ser	ttg Leu	gaa Glu	gaa Glu 435	Tyr	ctc Leu	aac Asn	aac Asn	gcc Ala 440	aag Lys	1348
	att Ile	tca Ser	ata Ile	tcg Ser 445	Ala	cct Pro	aca Thr	atc : Ile	ata Ile 450	Ser	cag Gln	ctt Leu	tat Tyr	ttt Phe 455	Thr	tta Leu	1396
	gca	aac	tcg	att	gat	gaa	aca	gċt	ato	gag	ago	ttg	tac	caa	tat	cat	1444

Ala	Asn	Ser 460	Ile	Asp	Glu	Thr	Ala 465	Ile	Glu	Ser	Leu	Tyr 470	Gln	Tyr	His	
aac Asn	ata Ile 475	ctt Leu	tac Tyr	cta Leu	tca Ser	gga Gly 480	acc Thr	ata Ile	tta Leu	agg Arg	ctt Leu 485	gct Ala	gac Asp	gat Asp	ctt Leu	1492
999 Gly 490	aca Thr	tca Ser	caa Gln	cat His	gag Glu 495	ctg Leu	gag Glu	aga Arg	gga Gly	gac Asp 500	gta Val	ccg Pro	aaa Lys	gca Ala	atc Ile 505	1540
cag Gln	tgc Cys	tac Tyr	atg Met	aat Asn 510	gac Asp	aca Thr	aat Asn	gct Ala	tcg Ser 515	gag Glu	aga Arg	gag Glu	gcg Ala	gtg Val 52 0	gaa Glu	1588
cac His	gtg Val	aag Lys	ttt Phe 525	ctg Leu	ata Ile	agg Arg	gag Glu	gcg Ala 530	tgg Trp	aag Lys	gag Glu	atg Met	aac Asn 535	acg Thr	gtc Val	1636
aca Thr	aca Thr	gcc Ala 540	agc Ser	gat As p	tgt Cys	ccg Pro	ttt Phe 545	acg Thr	gat Asp	gat Asp	ttg Leu	gtt Val 550	gcg Ala	gcc Ala	gca Ala	1684
gct Ala	aat Asn 555	ctt Leu	gca Ala	agg Arg	gcg Ala	gct Ala 560	cag Gln	ttt Phe	ata Ile	tat Tyr	ctc Leu 565	gac Asp	Gly ggg	gat Asp	G JA ā āā	1732
cat His 570	Gly Ggc	gtg Val	caa Gln	cac His	tca Ser 575	gaa Glu	ata Ile	cat His	caa Gln	cag Gln 580	atg Met	gga Gly	ggc Gly	ctg Leu	cta Leu 585	1780
		cct Pro		gtct	cgaat	caa a	atoga	aaaa	cc ca	aacci	tact	a tgi	tatc	cctc		1832
gati	aatat	at t	cttq	ggggt	tt aa	acat	gttt	a at	caaa	gttc	taa	ttdaa	aag	agct	gaatcg	1892
atc	ctca	aaa a	aaaa	1888	aa					•						1912
<210> 60 <211> 589 <212> PRT <213> Artificial Sequence																
	0> 6											_	_	_	•••	
Met 1	Ser	Ser	Ile	Ser 5	Ile	Asn	Ile	Ala	Met 10	Pro	Leu	Asn	Ser	Leu 15	H1S	
Asn	Phe	Glu	Arg 20	Lys	Pro	Ser	Lys	Ala 25	Trp	Ser	Thr	Ser	Cys 30	Thr	Ala	
Pro	Ala	Ala 35	Arg	Leu	Arg	Ala	Ser 40		Ser	Leu	Gln	Gln 45	Glu	Lys	Pro	

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 105 Val Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 120 Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 230 Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 250 255 Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365 Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu
405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 455 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly
465 470 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu
485 496 496

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr 580 585

<210> 61

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated (+)-sabinene synthase protein

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
sequence variant

<400> 61

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro

35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Ile Trp Asp Phe Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 170 Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 200 Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys

355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440 445

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 455 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly 465 470 475 480

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 505

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 535 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val 580 585

<210> 62

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated protein sequence variant

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
sequence variant

<400> 62 Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30 Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe Asn Tyr Val Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80 Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 105 Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 155 Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 200 Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 230 Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 375 380 Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425 430 Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 440 Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 520 Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 570 Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val

<210> 63 <211> 590 <212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

PCT/US98/20120 WO 99/15624

computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein sequence variant

<400> 63

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Glu Lys Pro

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe

Asn Tyr Ile Gln Ser Ile Asn Thr Pro Tyr Lys Glu Gln Arg His Phe

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 170

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 295 Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365 Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 390 Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp '/al Ser Leu Val Glu Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 425 Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 520 Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val

<210> 64 <211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
 sequence variant

<400> 64

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Val Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 135 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 215 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro

															240
225					230					235					_
Leu	His	Trp	Arg	Val 245	Gln	Gly	Leu	Glu	Ala 250	Arg	Trp	Phe	Leu	Asp 255	Ala
Tyr	Ala	Arg	Arg 260	Pro	Asp	Met	Asn	Pro 265	Leu	Ile	Phe	Lys	Leu 270	Ala	Lys
Leu	Asn	Phe 275	Asn	Ile	Val	Gln	Ala 280	Thr	Tyr	Gln	Glu	Glu 285	Leu	Lys	Asp
	290					293					Glu 300				
305					310					510					
Glu	Pro	His	Gln	Tyr 325	Ser	Tyr	Gln	Arg	1 2 ys	Met	Ala	Ala	Val	11e 335	Ile
Thr	Phe	Ile	Thr 340	Ile	Ile	Asp	Asp	Val 345	Туг	: Asp	Val	Tyr	Gly 350	Thr	Ile
		35	5				360	,							Lys
	370)				3/:	•								. Leu
385	•				390	J				55.	_				400
				40	>				7.1	•					1 Gl u 5
			42	0				42	5						o Ser
		43	5				44	U				• •	_		o Thr
	45	0				45	5					•			u Thr
Al 46	a Il 5	e G1	.u Se	r Le	и Ту 47	r Gl	n Ty	r Hi	.s As	n Il 47	e Le '5	u Ty	r Le	u Se	r Gly 480
Th	r Il	e L	eu Ar	g Le		a As	sp As	sp Le	eu Gi 41	Ly Th	r Se	r Gl	n Hi	.s G) 49	Lu Leu 95
Gl	u Aı	g G		sp Va 00	al P	co Ly	ys Al	La II	le G: 05	ln Cy	үз Ту	r Me	t As 51	n As LO	p Thr
As	in A	la S 5	er G] 15	Lu A	cg G	lu A	la Va 5	al G. 20	lu H	is Va	al Ly	/s Ph 52	ie Le 25	eu I	le Arg
G)	lu A	la T 30	rp L	ys G	Lu M	et A 5	sn Ti 35	hr V	al T	hr T	hr A	la Se 40	er A	вр С	ys Pro
Pi	ne T	hr A	sp A	sp L	eu V	al A	la A	la A	la A	la A	sn L	eu A	la A	rg A	la Ala

545 550 555 560

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val 580 585 590

<210> 65

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
 sequence variant

<400> 65

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His

1 5 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Leu Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 135 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn 165 170 175

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 215 Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 310 Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu-Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 505

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro 530 540

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val 580 585 590

<210> 66

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
 sequence variant

<400> 66

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 5 10

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro
35 40

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Val Glu Leu Ile Asp Asp Leu Gln Tyr 100 105 110

Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser 115 120 125

Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe 130 140

Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser 145 150 155 160

Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn

Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala 180 185 190

Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg 195 200 205

Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu 210 220

Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 225 230 235

Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala 245 250 255

Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys 260 265 270

Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 275 280 285

Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe 290 295 300

Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe 305 310 315 320

Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 325 330 335

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile 340 350

Glu Glu Leu Glu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365

Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu 370 380

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 395 400

Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 415

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser 420 425

Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 435 440

Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 450 460

Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly 465 470 475

Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 490 495

Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 500 510

Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 515 520 525

Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro

Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala 545 550 555

Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu 565 570 575

Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Prc Tyr Val 580 585

<210> 67

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(590)

<223> computer-generated (+)-sabinene synthase protein
sequence variant

<400> 67

Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 1 10 15

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Glu Lys Pro 35 40 45

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 60

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80

Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 95

Lys Met Glu Ala Ile Gln Gln Leu Glu Val Ile Asp Asp Leu Gln Tyr

			100					105					110		
Leu	Gly	Leu 115	Ser	Tyr	Phe	Phe	Gln 120	Asp	Glu	Ile	Lys	Gln 125	Ile	Leu	Ser
Ser	Ile 130	His	Asn	Glu	Pro	Arg 135	Tyr	Phe	His	Asn	Asn 140	Asp	Leu	Tyr	Phe
Thr 145	Ala	Leu	Gly	Phe	Arg 150	Ile	Leu	Arg	Gln	His 155	Gly	Phe	Asn	Val	Ser 160
Glu	Asp	Val	Phe	Asp 165	Cys	Phe	Lys	Ile	Glu 170	Lys	Cys	Ser	Asp	Phe 175	Asn
Ala	Asn	Leu	Ala 180	Gln	Asp	Thr	Lys	Gly 185	Met	Leu	Gln	Leu	Tyr 190	Glu	Ala
Ser	Phe	Leu 195	Leu	Arg	Glu	Gly	Glu 200	Asp	Thr	Leu	Glu	Leu 205	Ala	Arg	Arg
Phe	Ser 210	Thr	Arg	Ser	Leu	Arg 215	Glu	Lys	Phe	Asp	Glu 220	Gly	Gly	Asp	Glu
Ile 225	Asp	Glu	Asp	Leu	Ser 230	Ser	Trp	Ile	Arg	His 235	Ser	Leu	Asp	Leu	Pro 240
Leu	His	Trp	Arg	Val 245	Gln	Gly	Leu	Glu	Ala 250	Arg	Trp	Phe	Leu	Asp 255	Ala
Tyr	Ala	Arg	Arg 260	Pro	Asp	Met	Asn	Pro 265	Leu	Ile	Phe	Lys	Leu 270	Ala	Lys
Leu	Asn	Phe 275	Asn	Ile	Val	Gļn	Ala 280	Thr	Tyr	Gln	Glu	Glu 285	Leu	Lys	Asp
Ile	Ser 290		Trp	Trp	Asn	Ser 295	Ser	Cys	Leu	Ala	Glu 300	Lys	Leu	Pro	Phe
Val 305		Asp	Arg	, Ile	Val 310	Glu	Cys	Phe	Phe	Trp 315	Ala	Ile	Ala	Ala	Phe 320
Glu	Pro	His	s Glr	325	Ser	Туг	Gln	Arç	330 Lys	Met	Ala	Ala	Val	11e 335	Ile
The	Phe	e Ile	= Th:		e Ile	: Asp) Asp	Va) 345	L Tyr	: Asp	Val	Tyr	Gly 350	Thr	Ile
Glv	ı Glu	1 Let 35		ı Le	ı Let	Th	Asp 360		: Ile	a Arg	g Arg	365	Asp	Asn	Lys
Se	370	e Se	r Gl	n Lei	u Pro	37.	r Ty: 5	c Met	t Glr	ı Val	380	5 Туі)	Leu	ı Ala	Leu
Ту: 38:	r Ası 5	n Ph	e Va	l Se	r Gli 39	ı Ar	g Ala	а Ту	r Ası	39	e Lev 5	ı Ly:	s Asp	Glr	400
Ph	e As	n Se	r Il	e Pr 40	о Ту 5	r Le	u Gl	n Ar	g Se. 41	r Tr	p Val	l Se	r Le	1 Va:	l Glu 5

Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser

420 425 Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr 440 Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val <210> 68 <211> 590 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated protein sequence <220> <221> VARIANT <222> (1)..(590) <223> computer-generated (+)-sabinene synthase protein sequence variant <400> 68 Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala

Pro Ala Ala Arg Leu Arg Ala Ser Ser Ser Leu Gln Gln Glu Lys Pro

His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe

Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Val Gln Tyr 105 Leu Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 235 Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp 280 Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu

Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 385 390 Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 410 Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu 485 Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg 520 Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val <210> 69 <211> 590 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated protein sequence <220> <221> VARIANT <222> (1)..(590) <223> computer-generated(+)-sabinene synthase protein sequence <400> 69 Met Ser Ser Ile Ser Ile Asn Ile Ala Met Pro Leu Asn Ser Leu His 10

Asn Phe Glu Arg Lys Pro Ser Lys Ala Trp Ser Thr Ser Cys Thr Ala 20 25 30 Pro Ala Ala Arg Leu Arg Ala Ser Ser Leu Gln Gln Glu Lys Pro 35 40 His Gln Ile Arg Arg Ser Gly Asp Tyr Gln Pro Ser Leu Trp Asp Phe 50 55 Asn Tyr Ile Gln Ser Leu Asn Thr Pro Tyr Lys Glu Gln Arg His Phe 65 70 75 80 Asn Arg Gln Ala Glu Leu Ile Met Gln Val Arg Met Leu Leu Lys Val 85 90 Lys Met Glu Ala Ile Gln Gln Leu Glu Leu Ile Asp Asp Leu Gln Tyr Val Gly Leu Ser Tyr Phe Phe Gln Asp Glu Ile Lys Gln Ile Leu Ser Ser Ile His Asn Glu Pro Arg Tyr Phe His Asn Asn Asp Leu Tyr Phe Thr Ala Leu Gly Phe Arg Ile Leu Arg Gln His Gly Phe Asn Val Ser Glu Asp Val Phe Asp Cys Phe Lys Ile Glu Lys Cys Ser Asp Phe Asn Ala Asn Leu Ala Gln Asp Thr Lys Gly Met Leu Gln Leu Tyr Glu Ala Ser Phe Leu Leu Arg Glu Gly Glu Asp Thr Leu Glu Leu Ala Arg Arg Phe Ser Thr Arg Ser Leu Arg Glu Lys Phe Asp Glu Gly Gly Asp Glu Ile Asp Glu Asp Leu Ser Ser Trp Ile Arg His Ser Leu Asp Leu Pro 235 Leu His Trp Arg Val Gln Gly Leu Glu Ala Arg Trp Phe Leu Asp Ala Tyr Ala Arg Arg Pro Asp Met Asn Pro Leu Ile Phe Lys Leu Ala Lys Leu Asn Phe Asn Ile Val Gln Ala Thr Tyr Gln Glu Glu Leu Lys Asp Ile Ser Arg Trp Trp Asn Ser Ser Cys Leu Ala Glu Lys Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Phe Phe Trp Ala Ile Ala Ala Phe Glu Pro His Gln Tyr Ser Tyr Gln Arg Lys Met Ala Ala Val Ile Ile 330

Thr Phe Ile Thr Ile Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Ile Glu Glu Leu Glu Leu Leu Thr Asp Met Ile Arg Arg Trp Asp Asn Lys 355 360 365 Ser Ile Ser Gln Leu Pro Tyr Tyr Met Gln Val Cys Tyr Leu Ala Leu Tyr Asn Phe Val Ser Glu Arg Ala Tyr Asp Ile Leu Lys Asp Gln His 395 Phe Asn Ser Ile Pro Tyr Leu Gln Arg Ser Trp Val Ser Leu Val Glu 405 Gly Tyr Leu Lys Glu Ala Tyr Trp Tyr Tyr Asn Gly Tyr Lys Pro Ser Leu Glu Glu Tyr Leu Asn Asn Ala Lys Ile Ser Ile Ser Ala Pro Thr Ile Ile Ser Gln Leu Tyr Phe Thr Leu Ala Asn Ser Ile Asp Glu Thr 455 Ala Ile Glu Ser Leu Tyr Gln Tyr His Asn Ile Leu Tyr Leu Ser Gly 465 Thr Ile Leu Arg Leu Ala Asp Asp Leu Gly Thr Ser Gln His Glu Leu Glu Arg Gly Asp Val Pro Lys Ala Ile Gln Cys Tyr Met Asn Asp Thr 505 Asn Ala Ser Glu Arg Glu Ala Val Glu His Val Lys Phe Leu Ile Arg Glu Ala Trp Lys Glu Met Asn Thr Val Thr Thr Ala Ser Asp Cys Pro Phe Thr Asp Asp Leu Val Ala Ala Ala Ala Asn Leu Ala Arg Ala Ala Gln Phe Ile Tyr Leu Asp Gly Asp Gly His Gly Val Gln His Ser Glu Ile His Gln Gln Met Gly Gly Leu Leu Phe Gln Pro Tyr Val

<210> 70

<211> 1967

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (13)..(1785)
<223> computer-generated nucleic acid sequence encoding
 1,8-cineole synthase

gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat

Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His

20
25

cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc 147
Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys
30 40 45

tca cta caa atg ggt aat gag atc caa act gga cga cga act gga ggc 195 Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly

tac cag cct acc att tgg gat ttc agc acc att caa ttg ttc gac tct

Tyr Gln Pro Thr Ile Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser

70

75

gag tat aag gaa gag aag cac ttg atg agg gcc gca ggt atg ata gcc 291 Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala

Caa gtg aat atg ttg ttg cag gaa gaa gta gat tcg att caa cgg ttg 339
Gln Val Asn Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu
95 100

gag ttg att gat gac cta cga agg ctg ggt ata tct tgc cat ttt gac 38 Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp 110 125

ata gat gaa agt gat cta tac tca aca gcc ctt aga ttc aag ctc cta 483

Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu
145 150 155

aga caa tac gat ttt agc gtc tct caa gag gta ttt gat tgt ttc aag 531
Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys
160 165 170

aat gac aag ggt act gat ttc aag cca agc cta gtc gat gat act aga 579
Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg
175 180 185

gga ttg tta caa ttg tac gaa gct tcg ttt tta tca gca caa ggc gaa
Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu
190 200 205

gaa acc cta cat ctt gcc aga gat ttt gct act aaa ttt ctg cat aaa 675 Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys 210 215 220

					aaa Lys											723
					act Thr											771
					tat Tyr											819
					ttg Leu 275											867
					gcc Ala											915
					gtg Val											963
					gag Glu											1011
					gct Ala		-				-	•	-		-	1059
					gaa Glu 3 5 5											1107
					tca Ser											1155
					ttc Phe											1203
					ttc Phe											1251
					tca Ser										atg Met	1299
gga Gly 430	cat His	aaa Lys	cct Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atg Met	aag Lys 440	aat Asn	agt Ser	tgg Trp	ata Ile	tca Ser 445	1347
atc Ile	gga Gly	ggc Gly	atc Ile	CCC Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 455	ttt Phe	ttc Phe	cgg Arg	cta Leu	aca Thr 460	gat As p	1395

PCT/US98/20120 WO 99/15624

tcg att gag gaa gag gat gct gag agt atg cat aaa tac cat gat att Ser Ile Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile 465 470 475	3
gtt cgt gca tcg tgt act att cta agg ctt gct gat gat atg gga aca 1493 Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr 480 485 490	1
tcg ctg gat gag gtg gag aga ggc gac gtg ccc aaa tca gtt cag tgc 1539 Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys 495 500 505	•
tac atg aat gag aag aat gct tcg gaa gaa gaa gcg cga gag cat gtg Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val 510 525	,
cga tca ctc ata gac caa aca tgg aag atg atg aac aag gaa atg atg Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met 530 535 540	j
acg tca tca ttt tcc aaa tat ttt gta caa gtt tct gct aat ctt gca Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala 545 550 555	3
aga atg gcg caa tgg ata tac cag cat gaa tct gat gga ttt ggc atg Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met 560 565 570	L
caa cat toa ttg gtg aac aaa atg ctc aga ggg ttg ttg ttc gac cgc 1779 Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg 575 580 585	•
tat gag taactaatet tegeeegggt teeaaatgaa teaatetgtt gtgttgetgt 1839 Tyr Glu 590	5
tocacotgat atcaataata attagacaaa tgtttotgta ogggtggccc aacogtcagg 189	5
cccatttcgc tcatgttcat aataaataat aaaactgtta atcaataaca aaaaaaaaa 195	5
aaaaaaaaa aa 196	7
<210> 71 <211> 591 <212> PRT	

<212> PRT <213> Artificial Sequence

<400> 71

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro

55

Thr Ile Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 150 Asp Phe Ser Val Ser Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175 Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gin Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gin Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu

370 375 380 Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu 410 405 Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 520 Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 555 Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu

<210> 72

<211> 1967

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (13)..(1785)

<223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase

<400> 72

gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51
Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro

				1				5				1	.0			
gcc Ala	aaa Lys 15	att Ile	ttt Phe	cac His	aat Asn	aac Asn 20	tta Leu	ttc Phe	agc Ser	gtg Val	att Ile 25	tca Ser	aaa Lys	cga Arg	cat His	99
cgt Arg 30	ttc Phe	agt Ser	act Thr	aca Thr	atc Ile 35	acc Thr	act Thr	cgt Arg	ggt Gly	ggc Gly 40	agg Arg	tgg T <i>r</i> p	gca Ala	cat His	tgc Cys 45	147
tca Ser	cta Leu	caa Gln	atg Met	ggt Gly 50	aat Asn	gag Glu	atc Ile	caa Gln	act Thr 55	gga Gly	cga Arg	cga Arg	act Thr	gga Gly 60	ggc Gly	195
tac Tyr	cag Gln	cct Pro	acc Thr 65	ctt Leu	tgg Trp	gat Asp	ttc Phe	agc Ser 70	acc Thr	ctt Leu	caa Gln	ttg Leu	ttc Phe 75	gac Asp	tct Ser	243
gag Glu	tat Tyr	aag Lys 80	gaa Glu	gag Glu	aag Lys	cac His	ttg Leu 85	atg Met	agg Arg	gcc Ala	gca Ala	ggt Gly 90	atg Met	ata Ile	Ala	291
caa Gln	gtg Val 95	aat Asn	atg Met	ttg Leu	ttg Leu	cag Gln 100	gaa Glu	gaa Glu	gta Val	gat Asp	tcg Ser 105	att Ile	caa Gln	cgg Arg	ttg Leu	339
gag Glu 110	ttg Leu	att Ile	gat Asp	gac Asp	cta Leu 115	cga Arg	agg Arg	ctg Leu	ggt Gly	ata Ile 120	tct Ser	tgc Cys	cat His	ttt Phe	gac Asp 125	387
cgc Arg	gag Glu	atc Ile	gtt Val	gaa Glu 130	ata Ile	tta Leu	aac Asn	tca Ser	aaa Lys 135	tat Tyr	tat Tyr	acc Thr	aac Asn	aat Asn 140	gag Glu	435
ata Ile	gat Asp	gaa Glu	agt Ser 145	gat Asp	cta Leu	tac Tyr	tca Ser	aca Thr 150	gcc Ala	ctt Leu	aga Arg	ttc Phe	aag Lys 155	ctc Leu	cta Leu	483
aga Arg	caa Gln	tac Tyr 160	gat Asp	ttt Phe	agc Ser	gtc Val	tct Ser 165	caa Gln	gag Glu	gta Val	ttt Phe	gat Asp 170	tgt Cys	ttc Phe	aag Lys	531
aat Asn	gac Asp 175	aag Lys	ggt Gly	act Thr	gat A sp	ttc Phe 180	aag Lys	cca Pro	agc Ser	cta Leu	gtc Val 185	gat A sp	gat Asp	act Thr	aga Arg	579
gga Gly 190	ttg Leu	tta Leu	caa Gln	ttg Leu	tac Tyr 195	gaa Glu	gct Ala	tcg Ser	ttt Phe	tta Leu 200	tca Ser	gca Ala	caa Gln	Gly	gaa Glu 205	627
gaa Glu	acc Thr	cta Leu	cat His	ctt Leu 210	gcc Ala	aga Arg	gat Asp	ttt Phe	gct Ala 215	act Thr	aaa Lys	ttt Phe	ctg Leu	cat His 220	aaa Lys	675
aga Arg	gta Val	cta Leu	gtt Val 225	Asp	aaa Lys	gac Asp	att Ile	aat Asn 230	Leu	tta Leu	tca Ser	tca Ser	att Ile 235	gaa Glu	egt Arg	723
gcg Ala	ttg Leu	gag Glu	ttg Leu	cct Pro	act Thr	cat His	tgg Trp	agg Arg	gtt Val	caa Gln	atg Met	ccc Pro	aac Asn	gca Ala	aga Arg	771

		240					245					250				
													ecg Pro			819
cta Leu 270	gaa Glu	cta Leu	gct Ala	aaa Lys	ttg Leu 275	gac Asp	ttc Phe	aat Asn	atg Met	gtt Val 280	caa Gln	gca Ala	caa Gln	ttt Phe	caa Gln 285	867
													ggt Gly			915
													tac Tyr 315			963
													agg Arg			1011
													gtc Val			1059
													gct Ala			1107
													atg Met			1155
													tat Tyr 395			1203
													aaa Lys			1251
													tac Tyr			1299
													tgg Trp			1347
													cta Leu			1395
tcg Ser	att Ile	gag Glu	gaa Glu 465	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att. Ile	1443
gtt Val	cgt Arg	gca Ala	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp	atg Met	gga Gly	aca Thr	1491

		400					465					490				
							ggc Gly									1539
							tcg Ser									1587
							tgg Trp									1635
							ttt Phe									1683
							cag Gln 565									1731
							atg Met									1779
	gag Gl u	taad	taat	ct t	cgc	ccgg	gt to	caaa	atgaa	a tca	aatct	tgtt	gtgi	ttgct	:gt	1935
tcca	cct	gat a	tcaa	taat	a at	taga	caa	a tgi	tttci	tgta	cgg	gtgg	ccc a	aacc	gtcagg	1895
ccca	tttc	egc t	cat	jttca	at aa	ataaa	ataat	t aaa	aacto	gtta	atca	aata	aca a	aaaa	aaaaa	1955
aaaa	aaaa	aaa a	ıa													1967
<211 <212)> 73 l> 59 l> PF l> Ar)1 RT	i c ial	l Sed	quen	ce .										
	> 73		T	T1-	W	~ 1 ~	Val	V-1	T1 -	D	T	D	21-	T 140	Tla	
1	Jer	Ser	Deu	5	Mec	GIII	Val	Val	10	FIG	rys	FIG	Λια	15	110	
Phe	His	Asn	Asn 20	Leu	Phe	Ser	Val	Ile 25	Ser	Lys	Arg	His	Arg 30	Phe	Ser	
Thr	Thr	Ile 35	Thr	Thr	Arg	Gly	Gly 40	Arg	Trp	Ala	His	Cys 45	Ser	Leu	Gln	
Met	Gly 50	Asn	Glu	Ile	Gln	Thr 55	Glý	Arg	Arg	Thr	G1 y	Gly	Tyr	Gln	Pro	
Thr 65	Leu	Trp	Asp	Phe	Ser 70	Thr	Leu	Gln	Leu	Phe 75	Asp	Ser	Glu	Tyr	Lys 80	
G1 u	Glu	Lys	His	Leu 85		Arg	Ala	Ala	Gly 90		Ile	Ala	Gln	Val 95	Asn	

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 235 Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Glu Leu 280 Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu 405

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 515 520 Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu <210> 74 <211> 1967 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <222> (13)..(1785) <223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro

gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His

cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc

20

Arg 30	Phe	Ser	Thr	Thr	Ile 35	Thr	Thr	Arg	Gly	Gly 40	Arg	Trp	Ala	His	Cys 45	
tca Ser	cta Leu	caa Gln	atg Met	ggt Gly 50	aat Asn	gag Glu	atc Ile	caa Gln	act Thr 55	gga Gly	cga Arg	cga Arg	act Thr	gga Gly 60	ggc Gly	195
tac Tyr	cag Gln	cct Pro	acc Thr 65	ctt Leu	tgg Trp	gat Asp	ttc Phe	agc Ser 70	acc Thr	att Ile	caa Gln	gtg Val	ttc Phe 75	gac Asp	tct Ser	243
gag Glu	tat Tyr	aag Lys 80	gaa Glu	gag Glu	aag Lys	cac His	ttg Leu 85	atg Met	agg Arg	gcc Ala	gca Ala	ggt Gly 90	atg Met	ata Ile	gcc Ala	291
caa Gln	gtg Val 95	aat Asn	atg Met	ttg Leu	ttg Leu	cag Gln 100	gaa Glu	gaa Glu	gta Val	gat Asp	tcg Ser 105	att Ile	caa Gln	cgg Arg	ttg Leu	339
gag Glu 110	ttg Leu	att Ile	gat Asp	gac Asp	cta Leu 115	cga Arg	agg Arg	ctg Leu	ggt Gly	ata Ile 120	tct Ser	tgc Cys	cat His	ttt Phe	gac Asp 125	387
c gc Arg	gag Glu	atc Ile	gtt Val	gaa Glu 130	ata Ile	tta Leu	aac Asn	tca Ser	aaa Lys 135	tat Tyr	tat Tyr	acc Thr	aac Asn	aat Asn 140	gag Glu	435
ata Ile	gat Asp	gaa Glu	agt Ser 145	gat Asp	cta Leu	tac Tyr	tca Ser	aca Thr 150	gcc Ala	ctt Leu	aga Arg	ttc Phe	aag Lys 155	ctc Leu	cta Leu	483
aga Arg	caa Gln	tac Tyr 160	gat Asp	ttt Phe	agc Ser	gtc Val	tct Ser 165	caa Gln	gag Glu	gta Val	ttt Phe	gat Asp 170	tgt Cys	ttc Phe	aag Lys	531
aat Asn	gac Asp 175	aag Lys	ggt Gly	act Thr	gat Asp	ttc Phe 180	aag Lys	cca Pro	agc Ser	cta Leu	gtc Val 185	gat Asp	gat Asp	act Thr	aga Arg	579
gga Gly 190	ttg Leu	tta Leu	caa Gln	ttg Leu	tac Tyr 195	gaa Glu	gct Ala	tcg Ser	ttt Phe	tta Leu 200	tca Ser	gca Ala	caa Gln	Gly	gaa Glu 205	627
gaa Glu	acc Thr	cta Leu	cat His	ctt Leu 210	gcc Ala	aga Arg	gat Asp	ttt Phe	gct Ala 215	act Thr	aaa Lys	ttt Phe	ctg Leu	cat His 220	aaa Lys	675
aga Arg	gta Val	cta Leu	gtt Val 225	Asp	aaa Lys	gac Asp	att Ile	aat Asn 230	Leu	tta Leu	tca Ser	tca Ser	att Ile 235	gaa Glu	c gt Ar g	723
gcg Ala	ttg Leu	gag Glu 240	Leu	cct	act Thr	cat His	tgg Trp 245	Arg	gtt Val	caa Gln	atg Met	ccc Pro 250	Asn	gca Ala	aga Arg	7 71
tcc Ser	ttc Phe 255	att Ile	gat Asp	gct Ala	tat Tyr	aag Lys 260	Arg	aga Arg	ccc Pro	gac Asp	atg Met 265	Asn	ccg Pro	act Thr	gtg Val	8 19
cta	gaa	cta	gct	aaa	ttg	gac	ttc	aat	atg	gtt	caa	gca	caa	ttt	caa	867

Leu 270	Glu	Leu	Ala	Lys	Leu 275	Asp	Phe	Asn	Met	Val 280	Gln	Ala	Gln	Phe	Gln 285	
caa Gln	gag Glu	ctc Leu	aaa Lys	gag Glu 290	gcc Ala	tct Ser	a gg A rg	tgg Trp	tgg Trp 295	aat Asn	agt Ser	acg Thr	ggt Gly	ctt Leu 300	gtc Val	915
cac His	gag Glu	ctt Leu	ccc Pro 305	ttt Phe	gtg Val	aga Arg	gat Asp	agg Arg 310	att Ile	gtg Val	gaa Glu	tgc Cys	tac Tyr 315	tac Tyr	tg g Trp	963
acg Thr	aca Thr	gga Gly 320	gtg Val	gtt Val	gag Glu	cgt Arg	cgt Arg 325	gaa Glu	cat His	gga Gly	tac Tyr	gag Glu 330	agg Arg	ata Ile	atg Met	1011
ctc Leu	acc Thr 335	aaa Lys	ata Ile	aat Asn	gct Ala	ctt Leu 340	gtt Val	aca Thr	aca Thr	ata Ile	gac Asp 345	gat A sp	gtc Val	ttt Phe	gat Asp	1059
att Ile 350	tat Tyr	ggt Gly	acg Thr	ctt Leu	gaa Glu 355	gag Glu	cta Leu	caa Gln	cta Leu	ttc Phe 360	aca Thr	act Thr	gct Ala	att Ile	caa Gln 365	1107
aga Arg	tgg Trp	gat Asp	att Ile	gaa Glu 370	tca Ser	atg Met	aag Lys	caa Gln	ctc Leu 375	cct Pro	cct Pro	tac Tyr	atg Met	caa Gln 380	ata Ile	1155
tgt Cys	tat Tyr	ctt Leu	gct Ala 385	ctc Leu	ttc Phe	aac Asn	ttt Phe	gtg Val 39 0	aat Asn	gag Glu	atg Met	gct Ala	tat Tyr 395	gat Asp	act Thr	1203
ctt Leu	a gg A rg	gat Asp 400	aaa Lys	ggt Gly	ttc Phe	aac Asn	tcc Ser 405	acc Thr	cca Pro	tat Tyr	cta Leu	cga Arg 410	aaa Lys	gcg Ala	tgg Trp	1251
gtt Val	gat Asp 415	ttg Leu	gtt Val	gag Glu	tca Ser	tat Tyr 420	cta Leu	ata Ile	gag Glu	gca Ala	aag Lys 425	tgg Trp	tac Tyr	tac Tyr	atg Met	1299
gga Gly 430	cat His	aaa Lys	cct Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atg Met	aag Lys 440	aat Asn	agt Ser	tgg Trp	ata Ile	tca Ser 445	1347
atc Ile	gga Gly	ggc Gly	atc Ile	ecc Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 455	ttt Phe	ttc Phe	cgg Arg	cta Leu	aca Thr 460	gat Asp	1395
tcg Ser	att Ile	gag Glu	gaa Glu 465	Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att	1443
gtt Val	cgt Arg	gca Ala 480	Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	aca Thr	1491
tcg Ser	ctg Leu 495	gat Asp	gag Glu	gtg Val	gag Glu	aga Arg 500	Gly	gac Asp	gtg Val	ccc Pro	aaa Lys 505	tca Ser	gtt Val	cag Gln	tgc Cys	1539
tac	atg	aat	gag	aag	aat	gct	tcg	gaa	gaa	gaa	gcg	cga	gag	cat	gt g	1587

Tyr 510	Met	neA	Glu	Lys	Asn 515	Ala	Ser	Glu	Glu	Glu 520	Ala	Arg	Glu	His	Val 525	
cga Arg	tca Ser	ctc Leu	Ile	gac Asp 530	caa Gln	aca Thr	tgg Trp	aag Lys	atg Met 53 5	atg Met	aac Asn	aag Lys	gaa Glu	atg Met 54 0	atg Met	1635
acg Thr	tca Ser	tca Ser	ttt Phe 545	tcc Ser	aaa Lys	tat Tyr	ttt Phe	gta Val 550	caa Gln	gtt Val	tct Ser	gct Ala	aat Asn 555	ctt Leu	gca Ala	1683
aga Arg	atg Met	gcg Ala 560	caa Gln	tgg Trp	ata Ile	tac Tyr	cag Gln 565	cat His	gaa Glu	tct Ser	gat Asp	gga Gly 570	ttt Phe	ggc Gly	atg Met	1731
caa Gln	cat H1s 575	tca Ser	ttg Leu	gtg Val	aac Asn	aaa Lys 580	atg Met	ctc Leu	aga Arg	Gly ggg	ttg Leu 585	ttg Leu	ttc Phe	gac Asp	egc Arg	1779
tat Tyr 590		taac	taat	ct t	cgc	ecgg	gt to	caa	tgaa	a tca	atct	gtt	gtgt	tgci	igt	1835
teca	ecct	gat a	tcaa	taat	ta at	taga	acaaa	ı tgi	ttct	igta	cggg	tggo	cc a	acce	gtcagg	1895
															aaaaa	
aaa	aaaa	aa a	ıa													1967
<21:	0> 7: 1> 5: 2> P: 3> A:	91	icia	l Se	quen	ce										
<21: <21: <21: <40:	1> 5: 2> P: 3> A: 0> 7:	91 RT rtifi			-									•	T	
<21: <21: <21: <40:	1> 5: 2> P: 3> A: 0> 7:	91 RT rtifi			-		Val	Val	Ile	Pro	Lys	Pro	Ala	Lys 15	Ile	
<21: <21: <21: <40: Met	1> 5: 2> P: 3> A: 0> 7: Ser	91 RT rtifi 5 Ser	Leu	Ile 5	Met	Gln			.10	Pro Lys				15		
<21: <21: <21: <40: Met 1	1> 5: 2> P: 3> A: 0> 7: Ser	91 RT rtifi Ser Asn	Leu Asn 20	Ile 5 Leu	Met Phe	Gln Ser	Val	Ile 25	.10 Ser		Arg	His	Arg 30	15 Phe	Ser	
<21: <21: <21: <40: Met 1 Phe	1> 5: 2> P: 3> A: 0> 7: Ser His	91 RT rtifi 5 Ser Asn Ile 35	Leu Asn 20 Thr	Ile 5 Leu Thr	Met Phe Arg	Gln Ser Gly	Val Gly 40	Ile 25 Arg	.10 Ser Trp	Lys	Arg His	His Cys 45	Arg 30 Ser	15 Phe Leu	Ser Gln	
<21: <21: <21: <40: Met 1 Phe Thr	1> 5: 2> P: 3> A: 0> 7: Ser His Thr Gly 50	91 RT rtifi 5 Ser Asn Ile 35	Leu Asn 20 Thr	Ile 5 Leu Thr	Met Phe Arg	Gln Ser Gly Thr 55	Val Gly 40 Gly	Ile 25 Arg Arg	.10 Ser Trp Arg	Lys Ala	Arg His Gly 60	His Cys 45 Gly	Arg 30 Ser Tyr	15 Phe Leu Gln	Ser Gln Pro	
<21: <21: <21: <40: Met 1 Phe Thr Met Thr 65	1> 5: 2> P: 3> A: 0> 7: Ser His Thr Gly 50	91 RT rtiff Ser Asn Ile 35 Asn	Leu Asn 20 Thr Glu Asp	Ile 5 Leu Thr Ile	Met Phe Arg Gln Ser 70	Gln Ser Gly Thr 55	Val Gly 40 Gly	Ile 25 Arg Arg Gln	.10 Ser Trp Arg Val	Lys Ala Thr Phe 75 Met	Arg His Gly 60 Asp	His Cys 45 Gly Ser	Arg 30 Ser Tyr	15 Phe Leu Gln Tyr	Ser Gln Pro Lys 80 Asn	
<21: <21: <21: <400 Met 1 Phe Thr Met Thr 65 Glu	1> 52> PP 3> A 0> 7 Ser His Thr Gly 50 Leu	Pl RT rtiff Ser Asn Ile 35 Asn Trp	Leu Asn 20 Thr Glu Asp	Ile 5 Leu Thr Ile Phe Leu 85	Met Phe Arg Gln Ser 70	Gln Ser Gly Thr 55 Thr	Val Gly 40 Gly Ile	Ile 25 Arg Arg Gln	.10 Ser Trp Arg Val Gly 90	Lys Ala Thr Phe 75 Met	Arg His Gly 60 Asp	His Cys 45 Gly Ser	Arg 30 Ser Tyr Glu	Phe Leu Gln Tyr Val 95 Leu	Ser Gln Pro Lys 80 Asn	

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 360 Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gin Trp Ile Tyr Gin His Glu Ser Asp Gly Phe Gly Met Gin His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu <210> 76 <211> 1967 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <222> (13)..(1785) <223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase gatcaccaca ag atg tog agt ott ata atg caa gtt gtt att oot aag oca 51 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro gcc aaa att tit cac aat aac tta tic agc gig att tca aaa cga cat Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His 20 cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys 35 195 tca cta caa atg ggt aat gag atc caa act gga cga cga act gga ggc Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly

tac Tyr	cag Gln	cct Pro	acc Thr 65	ctt Leu	tgg Trp	gat Asp	ttc Phe	agc Ser 70	acc Thr	att Ile	caa Gln	ttg Leu	ttc Phe 75	gac Asp	tct Ser	243
gag Glu	tat Tyr	aag Lys 80	gaa Glu	gag Glu	aag Lys	cac His	gtg Val 85	atg Met	agg Arg	gcc Ala	gca Ala	ggt Gly 90	atg Met	ata Ile	gcc Ala	291
caa Gln	gtg Val 95	aat Asn	atg Met	ttg Leu	ttg Leu	cag Gln 100	gaa Glu	gaa Glu	gta Val	gat Asp	tcg Ser 105	att Ile	caa Gln	cgg Arg	ttg Leu	339
gag Glu 110	ttg Leu	att Ile	gat Asp	gac As p	cta Leu 115	cga Arg	agg Arg	ctg Leu	ggt Gly	ata Ile 120	tct Ser	tge Cys	cat His	ttt Phe	gac Asp 125	387
c gc A rg	gag Glu	atc Ile	gtt Val	gaa Glu 130	ata Ile	tta Leu	aac Asn	tca Ser	aaa Lys 135	tat Tyr	tat Tyr	acc Thr	aac Asn	aat Asn 140	gag Glu	435
ata Ile	gat Asp	gaa Glu	agt Ser 145	gat Asp	cta Leu	tac Tyr	tca Ser	aca Thr 150	gcc Ala	ctt Leu	aga Arg	ttc Phe	aag Lys 155	ctc Leu	cta Leu	483
Aṛg	Gln	Tyr 160	Asp	Phe	Ser	Val	Ser 165	Gln	gag Glu	Val	Phe	Asp 170	Cys	Phe	Lys	531
Asn	175	Lys	Gly	Thr	Asp	Phe 180	Lys	Pro	agc Ser	Leu	Val 185	Asp	Asp	Thr	Arg	579
190	Leu	Leu	Gln	Leu	Tyr 195	Glu	Ala	Ser	ttt Phe	Leu 200	Ser	Ala	Gln	Gly	Glu 205	627
GLu	Thr	Leu	His	Leu 210	Ala	Arg	Asp	Phe	gct Ala 215	Thr	Lys	Phe	Leu	His 220	Lys	675
Arg	Val	Leu	Val 225	Asp	Lys	Asp	Ile	Asn 230	ctc Leu	Leu	Ser	Ser	11e 235	Glu	Arg	723
Ala	Leu	Glu 240	Leu	Pro	Thr	His	Trp 245	Arg		Gln	Met	Pro 250	Asn	Ala	Arg	771
	255	lle	Asp	Ala	Tyr	Lys 260	Arg	Arg	Pro	Asp	Met 265	Asn	Pro	Thr	Val	819
270	GIU	Leu	Ala	Lys	Leu 275	Asp	Phe	Asn	atg Met	Val 280	Gln	Ala	Gln	Phe	Gln 285	867
Caa Gln	gag Glu	ctc Leu	Lys	gag Glu 2 9 0	gcc Ala	tct Ser	agg Arg	tgg Trp	tgg Trp 295	aat Asn	agt Ser	acg Thr	ggt Gly	ctt Leu 300	gtc Val	91 5

cac His	gag Glu	ctt Leu	ccc Pro 305	ttt Phe	gtg Val	aga Arg	gat Asp	agg Arg 310	att Ile	gtg Val	gaa Glu	tgc Cys	tac Tyr 315	tac Tyr	tgg Trp	963
acg Thr	aca Thr	gga Gly 320	gtg Val	gtt Val	gag Glu	cgt Arg	cgt Arg 325	gaa Glu	cat His	gga Gly	tac Tyr	gag Glu 330	agg Arg	ata Ile	atg Met	101Ý
								aca Thr								1059
att Ile 350	tat Tyr	ggt Gly	acg Thr	ctt Leu	gaa Glu 355	gag Glu	cta Leu	caa Gln	cta Leu	ttc Phe 360	aca Thr	act Thr	gct Ala	att Ile	caa Gln 365	1107
aga Arg	tgg Trp	gat Asp	att Ile	gaa Glu 370	tca Ser	atg Met	aag Lys	caa Gln	ctc Leu 375	cct Pro	cct Pro	tac Tyr	atg Met	caa Gln 380	ata Ile	1155
								gtg Val 390								1203
								acc Thr								1251
								ata Ile								1299
								tat Tyr								1347
atc Ile	gga Gly	ggc Gly	atc Ile	ccc Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 45 5	ttt Phe	ttc Phe	cgg Arg	cta Leu	aca Thr 460	gat As p	1395
tcg Ser	att Ile	gag Glu	gaa Glu 465	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att Ile	1443
								agg Arg								1491
tcg Ser	ctg Leu 495	gat Asp	gag Glu	gtg Val	gag Glu	aga Arg 500	Gly	gac Asp	gtg Val	ccc Pro	aaa Lys 505	tca Ser	gtt Val	cag Gln	tgc Cys	1539
tac Tyr 510	atg Met	aat Asn	gag Glu	aag Lys	aat Asn 515	gct Ala	tcg Ser	gaa Glu	gaa Glu	gaa Glu 520	gcg Ala	cga Arg	gag Glu	cat His	gtg Val 525	1587
								aag Lys								1635

Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala 545 550 555	1683
aga atg gcg caa tgg ata tac cag cat gaa tct gat gga ttt ggc atg Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met 560 565 570	1731 ⁹
caa cat tca ttg gtg aac aaa atg ctc aga ggg ttg ttg ttc gac cgc Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg 575 580 585	1779
tat gag taactaatct tcgcccgggt tccaaatgaa tcaatctgtt gtgttgctgt Tyr Glu 590	1835
tccacctgat atcaataata attagacaaa tgtttctgta cgggtggccc aaccgtcagg	1895
cccatttcgc tcatgttcat aataaataat aaaactgtta atcaataaca aaaaaaaaa	1955
aaaaaaaaa aa	1967

<210> 77 <211> 591 <212> PRT

<213> Artificial Sequence

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80

Glu Glu Lys His Val Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys

165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 245 250 255

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 260 265 270

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu 275 280 285

Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu 290 295 300

Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315

Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys 325 330 335

Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly
340 345 350

Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 355 360 365

Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 370 375 380

Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 390 395 400

Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu 405 410 415

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430

Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly
435 440 445

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu
450 455 460

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 465 470 475

Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp

				485					490					495		
Glu	Val	Glu	Arg 500	Gly	Asp	Val	Pro	Lys 505	Ser	Val	Gln	Суз	Tyr 510	Met	Asn	
Glu	Lys	Asn 515	Ala	Ser	Glu	Glu	Glu 520	Ala	Arg	Glu	His	Val 525	Arg	Ser	Leu	
Ile	Asp 530	Gln	Thr	Trp	Lys	Met 53 5	Met	Asn	Lys	Glu	Met 540	Met	Thr	Ser	Ser	
Phe 545	Ser	Lys	Tyr	Phe	Val 550	Gln	Val	Ser	Ala	Asn 555	Leu	Ala	Arg	Met	Ala 560	
Gln	Trp	Ile	Tyr.	Gln 565	His	Glu	Ser	Asp	Gly 570	Phe	Gly	Met	Gln	His 575	Ser	
Leu	Val	Asn	Lys 580	Met	Leu	Arg	Gly	Leu 585	Leu	Phe	Asp	Arg	Tyr 590	Glu		
<211 <212)> 78 .> 19 !> DN !> Ar	967 IA	.cial	. Sec	jueno	:e										
<220 <223	> De	escri	.ptic	on oi Jenei	Art	ific l nuc	cial cleid	Sequ aci	ence	: equer	ıce					
<220																
	> CI)S 13)	/170													
<223	\$> cc	mput	er-g	jene	ntha	i nuc	cleid	aci	id se	quer	nce e	encoc	iing			
<400	> 78	3					•									
gato	acca	ica a										il II			g cca ys Pro	51
gcc Ala	aaa Lys 15	att Ile	ttt Phe	cac His	aat Asn	aac Asn 20	tta Leu	ttc Phe	agc Ser	gtg Val	att Ile 25	tca Ser	aaa Lys	cga Arg	cat His	99
cgt Arg 30	ttc Phe	agt Ser	act Thr	aca Thr	atc Ile 35	acc Thr	act Thr	cgt Arg	ggt Gly	ggc Gly 40	agg Arg	tgg Trp	gca Ala	cat His	tgc Cys 45	147
tca Ser	cta Leu	caa Gln	atg Met	ggt Gly 50	aat Asn	gag Glu	atc Ile	caa Gln	act Thr 55	gga Gly	cga Arg	cga Arg	act Thr	gga Gly 60	ggc Gly	195
tac Tyr	cag Gln	cct Pro	acc Thr 65	ctt Leu	tgg Trp	gat As p	ttc Phe	agc Ser 70	acc Thr	att Ile	caa Gln	ttg Leu	ttc Phe 75	gac Asp	tct Ser	243
gag Glu	tat Tyr	aag Lys 80	gaa Glu	gag Glu	aag Lys	cac His	ttg Leu 85	atg Met	agg Arg	gcc Ala	gca Ala	ggt Gly 90	atg Met	cta Leu	gcc Ala	291

caa Gln	gtg Val 95	aat Asn	atg Met	ttg Leu	ttg Leu	cag Gln 100	gaa Glu	gaa Glu	gta Val	gat Asp	tcg Ser 105	att Ile	caa Gln	cgg Arg	ttg Leu	339
gag Glu 110	ttg Leu	att Ile	gat Asp	gac Asp	cta Leu 115	cga Arg	agg Arg	ctg Leu	ggt Gly	ata Ile 120	tct Ser	tgc Cys	cat His	ttt Phe	gac Asp 125	387
c gc Arg	gag Glu	atc Ile	gtt Val	gaa Glu 130	ata Ile	tta Leu	aac Asn	tca Ser	aaa Lys 135	tat Tyr	tat Tyr	acc Thr	aac Asn	aat Asn 140	gag Glu	435
ata Ile	gat Asp	gaa Glu	agt Ser 145	gat Asp	cta Leu	tac Tyr	tca Ser	aca Thr 150	gcc Ala	ctt Leu	aga Arg	ttc Phe	aag Lys 155	ctc Leu	cta Leu	483
aga Arg	caa Gln	tac Tyr 160	gat Asp	ttt Phe	agc Ser	gtc Val	tct Ser 165	caa Gln	gag Glu	gta Val	ttt Phe	gat Asp 170	tgt Cys	ttc Phe	aag Lys	531
aat Asn	gac Asp 175	aag Lys	ggt Gly	act Thr	gat Asp	ttc Phe 180	aag Lys	cca Pro	agc Ser	cta Leu	gtc Val 185	gat Asp	gat Asp	act Thr	aga Arg	579
gga Gly 190	ttg Leu	tta Leu	caa Gln	ttg Leu	tac Tyr 195	gaa Glu	gct Ala	tcg Ser	ttt Phe	tta Leu 200	tca Ser	gca Ala	caa Gln	ggc Gly	gaa Glu 205	627
gaa Glu	acc Thr	cta Leu	cat His	ctt Leu 210	gcc Ala	aga Arg	gat Asp	ttt Phe	gct Ala 215	act Th <i>r</i>	aaa Lys	ttt Phe	ctg Leu	cat His 220	aaa Lys	675
aga Arg	gta Val	cta Leu	gtt Val 225	gat Asp	aaa Lys	gac Asp	att Ile	aat Asn 230	ctc Leu	tta Leu	tca Ser	tca Ser	att Ile 235	gaa Glu	egt Arg	723
gcg Ala	ttg Leu	gag Glu 240	ttg Leu	cct Pro	act Thr	cat His	tgg Trp 245	agg Arg	gtt Val	caa Gln	atg Met	ccc Pro 250	aac Asn	gca Ala	aga Arg	771
tcc Ser	ttc Phe 255	att Ile	gat Asp	gct Ala	tat Tyr	aag Lys 260	agg Arg	a ga Arg	ccc Pro	gac Asp	atg Met 265	aat Asn	ccg Pro	act Thr	gtg Val	819
cta Leu 270	gaa Glu	cta Leu	gct Ala	aaa Lys	ttg Leu 275	gac Asp	ttc Phe	aat Asn	atg Met	gtt Val 280	caa Gln	gca Ala	caa Gln	ttt Phe	caa Gln 285	867
caa Gln	gag Glu	ctc Leu	aaa Lys	gag Glu 290	gcc Ala	tct Ser	agg Arg	tgg Trp	tgg Trp 295	aat Asn	agt Ser	acg Thr	ggt Gly	ctt Leu 300	gtc Val	915
					gtg Val				Ile							963
acg Thr	a ca Thr	gga Gly 32 0	Val	gtt Val	gag Glu	cgt Arg	cgt Arg 325	Glu	cat His	gga Gly	tac Tyr	gag Glu 330	Arg	ata Ile	atg Met	1011

										ata Ile						1059
att Ile 350	tat Tyr	ggt Gly	acg Thr	ctt Leu	gaa Glu 355	gag Glu	cta Leu	caa Gln	cta Leu	ttc Phe 360	aca Thr	act Thr	gct Ala	att Ile	caa Gln 365	1107
aga Arg	tgg Trp	gat Asp	att Ile	gaa Glu 370	tca Ser	atg Met	aag Lys	caa Gln	ctc Leu 375	cct Pro	cct Pro	tac Tyr	atg Met	caa Gln 380	ata Ile	1155
										gag Glu						1203
ctt Leu	agg Arg	gat Asp 400	aaa Lys	ggt Gly	ttc Phe	aac Asn	tcc Ser 405	acc Thr	cca Pro	tat Tyr	cta Leu	cga Arg 410	aaa Lys	gcg Ala	tg g Trp	1251
										gca Ala						1299
gga Gly 430	cat His	aaa Lys	cct Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atg Met	aag Lys 440	aat Asn	agt Ser	tgg Trp	ata Ile	tca Ser 445	1347
										ttt Phe						1395
tcg Ser	att Ile	gag Glu	gaa Glu 4 6 5	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att Ile	1443
gtt Val	cgt Arg	gca Ala 480	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	aca Thr	1491
										ccc						1539
										gaa Glu 520						1587
										atg Met						1635
acg Thr	tca Ser	tca Ser	ttt Phe 545	tcc Ser	aaa Lys	tat Tyr	ttt Phe	gta Val 550	caa Gln	gtt Val	tct Ser	gct Ala	aat Asn 555	ctt Leu	gca Ala	1683
aga Arg	atg Met	gcg Ala 560	caa Gln	tgg Trp	ata Ile	tac Tyr	cag Gln 565	cat His	gaa Glu	tct Ser	gat Asp	gga Gly 570	ttt Phe	ggc Gly	atg Met	1731

					aac Asn											1779
	tat gag taactaatet tegeeegggt tecaaatgaa teaatetgtt gtgttgetgt Tyr Glu 590														1835	
taca	tecaectgat atcaataata attagacaaa tgtttetgta egggtggeee aacegteagg															1895
CCC	cccatttcgc tcatgttcat aataaataat aaaactgtta atcaataaca aaaaaaaaaa														1955	
aacanaacaa aa														1967		
<210> 79 <211> 591 <212> PRT <213> Artificial Sequence																
	> 79		•	~ 1 -	N 4 - A -	-1 -	41- 1	11- 7	- 1 -		*	5		*	* 1.	
met 1	ser	ser	Leu	11e	Met	GIN	vaı	ΛSΤ	10	Pro	гÃа	Pro	Ala	15	TTE	
Phe	His	Asn	Asn 20	Leu	Phe	Ser	Vāl	11e 25	Ser	Lys	Arg	His	Arg 30	Phe	Ser	
Thr	Thr	Ile 35	Thr	Thr	Arg	Gly	Gly 40	Arg	Trp	Ala	His	Cys 45	Ser	Leu	Gln	
Met	Gly 50	Asn	Glu	Ile	Gln	Thr 55	Gly	Arg	Arg	Thr	Gly 60	Gly	Tyr	Gln	Pro	
Thr 65	Leu	Trp	Asp	Phe	Ser 70	Thr	Ile	Gln	Leu	Phe 75	Asp	Ser	Glu	Tyr	Lys 80	
Glu	Glu	Lys	His	Leu 85	Met	Arg	Ala	Ala	90 90	Met	Leu	Ala	Gln	Val 95	Asn	
Met	Leu	Leu	Gln 100	Glu	Glu	Val	Asp	Ser 105	Ile	Gln	Arg	Leu	Glu 110	Leu	Ile	
Ąsp	Asp	Leu 115	Arg	Arg	Leu	Gly	11e 120	Ser	Cās	His	Phe	Asp 125	Arg	Glu	Ile	
Val	Glu 130	Ile	Leu	Asn	Ser	Lys 135	Tyr	Tyr	Thr	Asn	Asn 140	Glu	Ile	Asp	Glu	
Ser 145	qeA	Leu	Tyr	Ser	Thr 150	Ala	Leu	Arg	Phe	Lys 155	Leu	Leu	Arg	Gln	Tyr 160	
Asp	Phe	Ser	Val	Ser 165	Gln	Glu	Val	Phe	Asp 170	Cys	Phe	Lys	Asn	Asp 175	Lys	
Gly	Thr	Asp	Phe 180	Lys	Pro	Ser	Leu	Val 185	Asp	Asp	Thr	Arg	Gly 190	Leu	Leu	
Gln	Leu	Tyr 195	Glu	Ala	Ser	Phe	Leu 200	Ser	Ala	Gln	Gly	Glu 205	Glu	Thr	Leu	

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 360 Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly 440 Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 470 Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 485 Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 520

- 170 -

Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 530 535 Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu <210> 80 <211> 1967 <212> DNA <213> Artificial Sequence <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <222> (13)..(1785) <223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase <400> 80 gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His 20 cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys 35 tca cta caa atg ggt aat gag atc caa act gga cga cga act gga ggc Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly tac cag cot acc ctt tgg gat ttc agc acc att caa ttg ttc gac tct Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser 243 gag tat aag gaa gag aag cac ttg atg agg gcc gca ggt atg ata gcc Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala 291 339 caa ctg aat atg ttg ttg cag gaa gaa gta gat tcg att caa cgg ttg Gln Leu Asn Met Leu Cln Glu Glu Val Asp Ser Ile Gln Arg Leu 100 gag ttg att gat gac cta cga agg ctg ggt ata tct tgc cat ttt gac Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp

110	115	120	125
		:	
Arg Glu Ile Val Glu	lle Leu Asn Ser	aaa tat tat acc aac aa Lys Tyr Tyr Thr Asn Asi 135	Glu
		gcc ctt aga ttc aag cto Ala Leu Arg Phe Lys Leo 155	
		gag gta ttt gat tgt ttc Glu Val Phe Asp Cys Pho 170	
		agc cta gtc gat gat act Ser Leu Val Asp Asp The 185	
gga ttg tta caa ttg Gly Leu Leu Gln Leu 190	tac gaa gct tcg t Tyr Glu Ala Ser I 195	ttt tta tca gca caa ggo Phe Leu Ser Ala Gln Gly 200	gaa 627 Glu 205
	Ala Arg Asp Phe A	got act aaa ttt ctg ca Ala Thr Lys Phe Leu Hi: 215 220	Lys
		ctc tta tca tca att gad Leu Leu Ser Ser Ile Glu 235	
		gtt caa atg ccc aac gca Val Gln Met Pro Asn Ala 250	
		occ gac atg aat ccg act Pro Asp Met Asn Pro Th 265	
		atg gtt caa gca caa tt Met Val Gln Ala Gln Pho 280	
	Ala Ser Arg Trp 1	tgg aat agt acg ggt ct Trp Asn Ser Thr Gly Le 295	val
cac gag ctt ccc ttt His Glu Leu Pro Phe 305	gtg aga gat agg a Val Arg Asp Arg 3 310	att gtg gaa tgc tac tac Ile Val Glu Cys Tyr Ty: 315	tgg 963
		cat gga tac gag agg at His Gly Tyr Glu Arg Ild 330	
ctc acc aaa ata aat Leu Thr Lys Ile Asr 335	get ett gtt aca a Ala Leu Val Thr 3 340	aca ata gac gat gtc tt Thr Ile Asp Asp Val Pho 345	gat 1059 Asp
		cta ttc aca act gct at Leu Phe Thr Thr Ala Il	

350					355					360					365	
													atg Met			1155
tgt Cys	tat Tyr	ctt Leu	gct Ala 385	ctc Leu	ttc Phe	aac Asn	ttt Phe	gtg Val 390	aat Asn	gag Glu	atg Met	gct Ala	tat Tyr 395	gat Asp	act Thr	1203
													aaa Lys			1251
													tac Tyr			1299
													tgg Trp			1347
													cta Leu			1395
													cat His 475			1443
													atg Met			1491
													gtt Val			1539
tac Tyr 510	atg Met	aat Asn	gag Glu	aag Lys	aat Asn 515	gct Ala	tcg Ser	gaa Glu	gaa Glu	gaa Glu 520	gcg Ala	cga Arg	gag Glu	cat His	gtg Val 525	1587
cga Arg	tca Ser	ctc Leu	ata Ile	gac Asp 530	caa Gln	aca Thr	tgg Trp	aag Lys	atg Met 53 5	atg Met	aac Asn	aag Lys	gaa Glu	atg Met 540	atg Met	1635
acg Thr	tca Ser	tca Ser	ttt Phe 545	tcc Ser	aaa Lys	tat Tyr	ttt Phe	gta Val 550	caa Gln	gtt Val	tct Ser	gct Ala	aat Asn 555	ctt Leu	gca Ala	1683
aga Arg	atg Met	gcg Ala 560	caa Gln	tgg Trp	ata Ile	tac Tyr	cag Gln 565	cat His	gaa Glu	tct Ser	gat Asp	gga Gly 570	ttt Phe	ggc Gly	atg Met	1731
caa Gln	cat His 5 75	tca Ser	ttg Leu	gtg Val	aac Asn	aaa Lys 580	atg Met	ctc Leu	aga Arg	ggg Gly	ttg Leu 585	ttg Leu	ttc Phe	gac Asp	egc Arg	1779
tat Tyr		taad	taat	ict t	cgc	ccgg	gt to	caaa	atgaa	tca	aatct	gtt	gtgt	tge	tgt	1835

'590

tccacctgat atcaataata attagacaaa tgtttctgta cgggtggccc aaccgtcagg 1895 cccatttcgc tcatgttcat aataaataat aaaactgtta atcaataaca aaaaaaaaa 1955* aaaaaaaaaa aa 1967

<210> 81

<211> 591

<212> PRT

<213> Artificial Sequence

<400> 81

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40

Met Gly Asn Glu Ile Gln Thr Gly Arg Thr Gly Gly Tyr Gln Pro

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys 65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Leu Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190

Gin Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly 440 Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu
580 585 590

<210> 82

<211> 1967

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
 computer-generated nucleic acid sequence

<220>

<221> CDS

<222> (13)..(1785)

<223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase

<400> 82

gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro

10

gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat 99
Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His
15 20 25

cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc 147
Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys
30 40 45

tca cta caa atg ggt aat gag atc caa act gga cga cga act gga ggc 195 Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly 50 55 60

tac cag cct acc ctt tgg gat ttc agc acc att caa ttg ttc gac tct 243
Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser
75

gag tat aag gaa gag aag cac ttg atg agg gcc gca ggt atg ata gcc 291 Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala 80 85

Caa gtg aat atg gtg ttg cag gaa gaa gta gat tcg att caa cgg ttg 339 Gln Val Asn Met Val Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu 95 100

gag ttg att gat gac cta cga agg ctg ggt ata tct tgc cat ttt gac 387 Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp 110 125

ata gat gaa agt gat cta tac tca aca gcc ctt aga ttc aag ctc cta 483

Ile	qeA	Glu	Ser 145	Asp	Leu	Tyr	Ser	Thr 150	Ala	Leu	Arg	Phe	Lys 155	Leu	Leu	
							tct Ser 165									531
aat Asn	gac Asp 175	aag Lys	ggt Gly	act Thr	gat Asp	ttc Phe 180	aag Lys	cca Pro	agc Ser	cta Leu	gtc Val 185	gat Asp	gat Asp	act Thr	a ga A rg	579
							gct Ala									627
							gat Asp									675
							att Ile									723
							tgg Trp 245									771
							agg Arg									819
							ttc Phe									867
caa Gln	gag Glu	ctc Leu	aaa Lys	gag Glu 290	gcc Ala	tct Ser	agg Arg	tgg Trp	tgg Trp 295	aat Asn	agt Ser	acg Thr	ggt Gly	ctt Leu 300	gtc Val	915
cac His	gag Glu	ctt Leu	ccc Pro 305	ttt Phe	gtg Val	aga Arg	gat A sp	agg Arg 310	att Ile	gtg Val	gaa Glu	tgc Cys	tac Tyr 315	tac Tyr	tg g Trp	963
							cgt Arg 325									1011
							gtt Val									1059
							cta Leu									1107
aga Arg	tgg Trp	gat Asp	att Ile	gaa Glu 370	tca Ser	atg Met	a ag Lys	caa Gln	ctc Leu 375	cct Pro	cct Pro	tac Tyr	atg Met	caa Gln 380	ata Ile	1155
tgt	tat	ctt	gct	ctc	ttc	aac	ttt	gtg	aat	gag	atg	gct	tat	gat	act	1203

Cys	Tyr	Leu	Ala 385	Leu	Phe	Asn	Phe	Val 390	Asn	Glu	Met	Ala	Tyr 395	qeA	Thr	
ctt Leu	agg Arg	gat Asp 400	aaa Lys	ggt Gly	ttc Phe	aac Asn	tcc Ser 405	acc Thr	cca Pro	tat Tyr	cta Leu	cga Arg 410	aaa Lys	gcg Ala	tgg Trp	1251
gtt Val	gat Asp 415	ttg Leu	gtt Val	gag Glu	tca Ser	tat Tyr 420	cta Leu	ata Ile	gag Glu	gca Ala	aag Lys 425	tgg Trp	tac Tyr	tac Tyr	atg Met	1299
gga Gly 43 0	cat His	aaa Lys	cct Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atg Met	aag Lys 440	aat Asn	agt Ser	tgg Trp	ata Ile	tca Ser 445	1347
atc Ile	gga Gly	ggc Gly	atc Ile	ccc Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 455	ttt Phe	ttc Phe	cgg Arg	cta Leu	aca Thr 460	gat Asp	1395
tcg Ser	att Ile	gag Glu	gaa Glu 46 5	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att Ile	1443
gtt Val	cgt Arg	gca Ala 480	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	aca Thr	1491
tcg Ser	ctg Leu 495	gat Asp	gag Glu	gtg Val	gag Glu	aga Arg 500	ggc Gly	gac Asp	gtg Val	ccc Pro	aaa Lys 505	tca Ser	gtt Val	cag Gln	tgc Cys	1539
tac Tyr 510	atg Met	aat Asn	gag Glu	aag Lys	aat Asn 515	gct Ala	tcg Ser	gaa Glu	gaa Glu	gaa Glu 520	gcg Ala	cga Arg	gag Glu	cat His	gtg Val 525	1587
cga Arg	tca Ser	ctc Leu	ata Ile	gac Asp 530	caa Gln	aca Thr	tgg Trp	aag Lys	atg Met 535	atg Met	aac Asn	aag Lys	gaa Glu	atg Met 540	atg Met	1635
acg Thr	tca Ser	tca Ser	ttt Phe 545	tcc Ser	aaa Lys	tat Tyr	ttt Phe	gta Val 550	caa Gln	gtt Val	tct Ser	gct Ala	aat Asn 555	ctt Leu	gca Ala	1683
aga Arg	atg Met	gcg Ala 560	caa Gln	tgg Trp	ata Ile	tac Tyr	cag Gln 56 5	cat His	gaa Glu	tct Ser	gat Asp	gga Gly 570	ttt Phe	ggc Gly	atg Met	1731
caa Gln	cat His 575	tca Ser	ttg Leu	gtg Val	aac Asn	aaa Lys 580	atg Met	ctc Leu	aga Arg	ggg	ttg Leu 585	ttg Leu	ttc Phe	gac Asp	cgc Arg	1779
	Glu		ctaa	tct	tege	cc g g	gt t	ccaa	atga	a tc	aatc	tgtt	gtg	ttgc	t gt	1835
tec	acct	gat	atca	ataa	ta a	ttag	acaa	a tg	tttc	tgta	cgg	gtgg	ccc .	aacc	gtcagg	1895
ccc	attt	cgc	tcat	gttc	at a	ataa	ataa	t aa	aact	gtta	atc	aata	aca	aaaa	aaaaa	1955
aaa	aaaa	aaa	aa				•									1967

<210> 83 <211> 591 <212> PRT

<213> Artificial Sequence

<400> 83

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Val Leu Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 245 250 255

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 260 265 270

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu

285

Lys	Glu 290	Ala	Ser	Arg	Trp	Trp 295	Asn	Ser	Thr	Gly	Leu 300	Val	His	Glu	Leu

280

275

Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315

Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys 325 330 335

Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly 340 345 350

Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 355 360 365

Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 370 380

Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 390 395 400

Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu
405 410 415

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430

Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly
435 440 445

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu 450 455 460

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 465 470 480

Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 485 490 495

Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn 500 505

Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 515 520 525

Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 530 540

Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 545 550 555

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 580 585

<210> 84 <211> 1967 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated nucleic acid sequence <220> <221> CDS <222> (13)..(1785) <223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys tca cta caa atg ggt aat gag atc caa act gga cga act gga ggc Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly tac cag cot acc off tgg gat the age acc att caa ttg the gad tot Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser gag tat aag gaa gag aag cac ttg atg agg gcc gca ggt atg ata gcc Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala 85 339 caa gtg aat atg ttg ttg cag gaa gaa cta gat tcg att caa cgg ttg Gln Val Asn Met Leu Leu Gln Glu Glu Leu Asp Ser Ile Gln Arg Leu gag ttg att gat gac cta cga agg ctg ggt ata tet tgc cat ttt gac Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp 435 Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu 483 ata gat gaa agt gat cta tac tca aca gcc ctt aga ttc aag ctc cta Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu 150 aga caa tac gat ttt agc gtc tct caa gag gta ttt gat tgt ttc aag Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys 165

aat Asn	gac Asp 175	aag Lys	ggt Gly	act Thr	gat Asp	ttc Phe 180	aag Lys	cca Pro	agc Ser	cta Leu	gtc Val 185	gat Asp	gat Asp	act Th <i>r</i>	aga Arg	579 ∛
						gaa Glu										627
gaa Glu	acc Thr	cta Leu	cat His	ctt Leu 210	gcc Ala	aga Arg	gat Asp	ttt Phe	gct Ala 215	act Thr	aaa Lys	ttt Phe	ctg Leu	cat His 220	aaa Lys	675
						gac Asp										723
gcg Ala	ttg Leu	gag Glu 240	ttg Leu	cct Pro	act Thr	cat His	tgg Trp 245	agg Arg	gtt Val	caa Gln	atg Met	Pro 250	aac Asn	gca Ala	aga Arg	771
tcc Ser	ttc Phe 255	att Ile	gat Asp	gct Ala	tat Tyr	aag Lys 260	agg Arg	aga Arg	ccc Pro	gac As p	atg Met 265	aat Asn	ecg Pro	act Thr	gtg Val	819
						gac Asp										867
						tct Ser										915
cac His	gag Glu	ctt Leu	ccc Pro 305	ttt Phe	gtg Val	aga Arg	gat Asp	agg Arg 310	att Ile	gtg Val	gaa Glu	tgc Cys	tac Tyr 315	tac Tyr	tg g Tr p	963
acg Thr	aca Thr	gga Gly 320	gtg Val	gtt Val	gag Glu	cgt Arg	cgt Arg 325	gaa Glu	cat His	gga Gly	tac Tyr	gag Glu 330	agg Arg	ata Ile	atg Met	1011
						ctt Leu 340										1059
att Ile 350	tat Tyr	ggt Gly	acg Thr	ctt Leu	gaa Glu 355	gag Glu	cta Leu	caa Gln	cta Leu	ttc Phe 360	aca Thr	act Thr	gct Ala	att Ile	caa Gln 365	1107
a ga Arg	tgg Trp	gat Asp	att Ile	gaa Glu 370	tca Ser	atg Met	aag Lys	caa Gln	ctc Leu 375	cct Pro	cct Pro	tac Tyr	atg Met	caa Gln 380	ata Ile	1155
tgt Cys	tat Tyr	ctt Leu	gct Ala 385	ctc Leu	ttc Phe	aac Asn	ttt Phe	gtg Val 390	aat Asn	gag Glu	atg Met	gct Ala	tat Tyr 395	gat Asp	act Thr	1203
ctt Leu	agg Arg	gat Asp 400	aaa Lys	ggt Gly	ttc Phe	aac Asn	tcc Ser 405	acc Thr	cca Pro	tat Tyr	cta Leu	cga Arg 410	aaa Lys	gcg Ala	tgg Trp	1251

gtt Val	gat Asp 415	ttg Leu	gtt Val	gag Glu	tca Ser	tat Tyr 420	cta Leu	ata Ile	gag Glu	gca Ala	aag Lys 425	tgg Trp	tac Tyr	tac Tyr	atg Met	1299
gga Gly 430	cat His	aaa Lys	cct Pro	agt Ser	ttg Leu 435	gaa Glu	gaa Glu	tat Tyr	atç Met	aag Lys 440	aat Asn	agt Ser	tgg Trp	ata Ile	tca Ser 445	1347
atc Ile	gga Gly	ggc Gly	atc Ile	ccc Pro 450	att Ile	cta Leu	tct Ser	cat His	cta Leu 455	ttt Phe	ttc Phe	egg Arg	cta Leu	aca Thr 460	gat Asp	1395
tcg Ser	att Ile	g ag Glu	gaa Glu 465	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att Ile	1443
gtt Val	cgt Arg	gca Ala 480	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	aca Thr	1491
tcg Ser	ctg Leu 495	gat Asp	gag Glu	gtg Val	gag Glu	aga Arg 500	ggc Gly	gac As p	gtg Val	ccc Pro	aaa Lys 505	tca Ser	gtt Val	cag Gln	tgc Cys	1539
tac Tyr 510	atg Met	aat Asn	gag Glu	aag Lys	aat Asn 51 5	gct Ala	tcg Ser	gaa Glu	gaa Glu	gaa Glu 52 0	gcg Ala	cga Arg	gag Glu	cat His	gtg Val 525	1587
					caa Gln											1635
acg Thr	tca Ser	tca Ser	ttt Phe 545	tcc Ser	aaa Lys	tat Tyr	ttt Phe	gta Val 550	caa Gln	gtt Val	tct Ser	gct Ala	aat Asn 555	ctt Leu	gca Ala	1683
a ga Ar g	atg Met	gcg Ala 560	caa Gln	tgg Trp	ata Ile	tac Tyr	cag Gln 565	cat His	gaa Glu	tct Ser	gat A sp	gga Gly 570	ttt Phe	ggc Gly	atg Met	1731
caa Gln	cat His 575	tca Ser	ttg Leu	gtg Val	aac Asn	aaa Lys 580	atg Met	ctc Leu	aga Arg	G] À Gàà	ttg Leu 585	ttg Leu	ttc Phe	gac Asp	cgc Arg	1779 [.]
tat Tyr 590	gag Glu	taac	taat	ct t	cgc	cgg	gt to	caaa	tgaa	tca	atct	gtt	gtgt	tgct	gt	1835
tcca	ccto	jat a	tcas	taat	a at	taga	caaa	a tgt	ttc	gta	cgg	gtggd	cc a	acco	gtcagg	1895
CCC	ttt	gc t	cato	gttca	at aa	taaa	taat	: aaa	acto	tta	atca	ataa	ica a	aaaa	aaaaa	1955
aaaa	aaaa	aa a	ıa													1967

<210> 85 <211> 591 <212> PRT <213> Artificial Sequence

<400> 85 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10 15 Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30 Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45 Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95 Met Leu Leu Gln Glu Glu Leu Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 150 Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190 Gin Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu 290 295 300 Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly

Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 520 Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu

<210> 86

<211> 1967

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

computer-generated sequence

<220> <221> CDS <222> (13)..(1785) <223> computer-generated nucleic acid sequence encoding 1,8 cineole synthase <400> 86 gatcaccaca ag atg tcg agt ctt ata atg caa gtt gtt att cct aag cca 51 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro gcc aaa att ttt cac aat aac tta ttc agc gtg att tca aaa cga cat Ala Lys Ile Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His cgt ttc agt act aca atc acc act cgt ggt ggc agg tgg gca cat tgc Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys 35 tca cta caa atg ggt aat gag atc caa act gga cga cga act gga ggc 195 Ser Leu Gln Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly tac cag cot acc off tgg gat the age acc att caa ttg the gad tot Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser gag tat aag gaa gag aag cac ttg atg agg gcc gca ggt atg ata gcc Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala 291 caa gtg aat atg ttg ttg cag gaa gaa gta gat tcg ctt caa cgg ttg Gin Val Asn Met Leu Leu Gin Glu Glu Val Asp Ser Leu Gin Arg Leu 100 gag ttg att gat gac cta cga agg ctg ggt ata tct tgc cat ttt gac Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp 110 435 Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu 130 ata gat gaa agt gat cta tac tca aca gcc ctt aga ttc aag ctc cta 483 Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu 150 aga caa tac gat tit age gie tet caa gag gia tit gat tgt tie aag 531 Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys aat gac aag ggt act gat tto aag coa ago ota gto gat gat act aga 579 Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg gga ttg tta caa ttg tac gaa gct tcg ttt tta tca gca caa ggc gaa 627 Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu

				ctt Leu 210												675
aga Arg	gta Val	cta Leu	gtt Val 225	gat Asp	aaa Lys	gac Asp	att Ile	aat Asn 230	ctc Leu	tta Leu	tca Ser	tca Ser	att Ile 235	gaa Glu	egt Arg	723
				cct Pro												771
				gct Ala												819
				aaa Lys												867
				gag Glu 290												915
				ttt Phe												963
				gtt Val												1011
				aat Asn												1059
				ctt Leu												1107
				gaa Glu 370												1155
tgt Cys	tat Tyr	ctt Leu	gct Ala 385	ctc Leu	ttc Phe	aac Asn	ttt Phe	gtg Val 390	aat Asn	gag Glu	atg Met	gct Ala	tat Tyr 395	gat Asp	act Thr	1203
ctt Leu	agg Arg	gat Asp 400	aaa Lys	ggt Gly	ttc Phe	aac Asn	tcc Ser 405	acc Thr	cca Pro	tat Tyr	cta Leu	cga Arg 410	aaa Lys	gcg Ala	tgg Trp	1251
gtt Val	gat Asp 415	ttg Leu	gtt Val	gag Glu	tca Ser	tat Tyr 420	cta Leu	ata Ile	gag Glu	gca Ala	aag Lys 425	tgg Trp	tac Tyr	tac Tyr	atg M e t	1299
				agt Ser												1347

PCT/US98/20120 WO 99/15624

					att Ile											1395
tcg Ser	att Ile	gag Glu	gaa Glu 465	gag Glu	gat Asp	gct Ala	gag Glu	agt Ser 470	atg Met	cat His	aaa Lys	tac Tyr	cat His 475	gat Asp	att Ile	1443
gtt Val	egt Arg	gca Ala 480	tcg Ser	tgt Cys	act Thr	att Ile	cta Leu 485	agg Arg	ctt Leu	gct Ala	gat Asp	gat Asp 490	atg Met	gga Gly	aca Thr	1491
					gag Glu											1539
tac Tyr 510	atg Met	aat Asn	gag Glu	aag Lys	aat Asn 515	gct Ala	tcg Ser	gaa Glu	gaa Glu	gaa Glu 52 0	gcg Ala	cga Arg	gag Glu	cat His	gtg Val 525	1587
	Ser				caa Gln											1635
					aaa Lys											1683
					ata Ile											1731
					aac Asn											1779
tat Tyr 590		taad	taat	et 1	ege	eegge	jt to	caaa	tgaa	a tca	atc	gtt	gtgi	ttget	igt	1835
teca	cctç	gat a	tcaa	taa	a at	taga	caaa	a tgt	ttct	gta	cgg	gtggd	ccç a	aacco	gtcagg	1895
. ccca	ittt	ege t	cato	gttca	at aa	ataaa	taat	aaa	acto	gtta	atca	aataa	aca a	aaaa	148888	1955
8888	aaaa	aa a	ıa													1967

<210> 87 <211> 591 <212> PRT

<213> Artificial Sequence

<400> 87

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 10

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80 Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95 Met Leu Leu Gln Glu Glu Val Asp Ser Leu Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160 Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175 Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gin Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205 His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly 345

Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 520 Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu

- <210> 88
- <211> 591
- <212> PRT
- <213> Artificial Sequence
- <220>
- <223> Description of Artificial Sequence: computer-generated protein sequence
- <220>
- <221> VARIANT
- <222> (1)..(591)
- <223> computer-generated 1,8 cineole synthase protein sequence variant

<400> 88 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 15 Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55 60 Thr Ile Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys 65 70 75 80 Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125 Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly

PCT/US98/20120 WO 99/15624

305 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 375 Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 535 Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu <210> 89 <211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence variant

<400> 89

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55 60

Thr Leu Trp Asp Phe Ser Thr Leu Gln Leu Phe Asp Ser Glu Tyr Lys 65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 245 250 255

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 260 265 270

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 360 Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 390 395 400 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu 455 Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 470 Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 535 Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 585

<210> 90 <211> 591 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated protein sequence <220> <221> VARIANT <222> (1)..(591) <223> computer-generated 1,8 cineole synthase protein sequence variant <400> 90 Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile
1 5 10 15 Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser Thr Thr Ile Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 60 The Leu Trp Asp Phe Ser Thr Ile Gln Val Phe Asp Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95 Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu

215

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 230 Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 310 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly 345 Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 375 Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu 455 Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser

Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 545 550 555

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 580 595

<210> 91

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence variant

<400> 91

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
70 75 80

Glu Glu Lys His Val Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu

180 185 190

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 245 250 255

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 260 265 270

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu 275 280 285

Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu 290 295 300

Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315

Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys 325 330 335

Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly 340 345

Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 355 360 365

Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 370 375 380

Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 390 395 400

Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu
405 410 415

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430

Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly
435 440 445

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu
450 455 460

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 465 470 475 480

Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 495

Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn

500 505

Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 515 520 525

Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 530 540

Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 545 550 555 560

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu
580 585 590

<210> 92

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence variant

<400> 92

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser .20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln
35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55 60

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Leu Ala Gln Val Asn 90 95

Met Leu Leu Glu Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 185 Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 200 His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile 250 Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys 325 330 335 Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 375 Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu 455

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn 505 Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 550 Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 570 Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu <210> 93 <211> 591 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: computer-generated protein sequence <220> <221> VARIANT <222> (1)..(591) <223> computer-generated 1,8 cineole synthase protein

<400> 93

sequence variant

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln
35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro 50 55 60

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75 80

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Leu Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu 410

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430

Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly 435 440 445

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu 450 455 460

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 465 470 475

Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 490 495

Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn 500 505

Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu
515 525

Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 530 540

Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 545 550 555

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 575 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 580 595

<210> 94

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence

<400> 94

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 5 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro

50	55	60
----	----	----

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn Met Val Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile **Val** Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 360 Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu

370 375 380

Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 390 395

Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu 405 410 415

Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430

Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly
435 440 445

Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu
450 455 460

Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala 465 470 475 480

Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp 485 490 495

Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn 500 510

Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu 515 520 525

Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 530 540

Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 545 550 555 560

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 570 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu
585 590

<210> 95

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence

<400> 95

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1. 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln
35 40 45 Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn Met Leu Leu Gln Glu Glu Leu Asp Ser Ile Gln Arg Leu Glu Leu Ile Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 215 Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu 265 Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu 295 Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys

Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser 535 Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala 550 Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 580

<210> 96

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence

<400> 96

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Leu Gln Arg Leu Glu Leu Ile

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 235

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu

Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu 280

Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu 295 Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly 305 310 315 Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp 360 Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 390 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 425 Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser 565 Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 585

<210> 97 <211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: computer-generated protein sequence

<220>

<221> VARIANT

<222> (1)..(591)

<223> computer-generated 1,8 cineole synthase protein sequence variant

<400> 97

Met Ser Ser Leu Ile Met Gln Val Val Ile Pro Lys Pro Ala Lys Ile 1 10 15

Phe His Asn Asn Leu Phe Ser Val Ile Ser Lys Arg His Arg Phe Ser 20 25 30

Thr Thr Ile Thr Thr Arg Gly Gly Arg Trp Ala His Cys Ser Leu Gln 35 40 45

Met Gly Asn Glu Ile Gln Thr Gly Arg Arg Thr Gly Gly Tyr Gln Pro
50 60

Thr Leu Trp Asp Phe Ser Thr Ile Gln Leu Phe Asp Ser Glu Tyr Lys
65 70 75

Glu Glu Lys His Leu Met Arg Ala Ala Gly Met Ile Ala Gln Val Asn 85 90 95

Met Leu Leu Gln Glu Glu Val Asp Ser Ile Gln Arg Val Glu Leu Ile 100 105 110

Asp Asp Leu Arg Arg Leu Gly Ile Ser Cys His Phe Asp Arg Glu Ile 115 120 125

Val Glu Ile Leu Asn Ser Lys Tyr Tyr Thr Asn Asn Glu Ile Asp Glu 130 135 140

Ser Asp Leu Tyr Ser Thr Ala Leu Arg Phe Lys Leu Leu Arg Gln Tyr 145 150 155 160

Asp Phe Ser Val Ser Gln Glu Val Phe Asp Cys Phe Lys Asn Asp Lys 165 170 175

Gly Thr Asp Phe Lys Pro Ser Leu Val Asp Asp Thr Arg Gly Leu Leu 180 185 190

Gln Leu Tyr Glu Ala Ser Phe Leu Ser Ala Gln Gly Glu Glu Thr Leu 195 200 205

His Leu Ala Arg Asp Phe Ala Thr Lys Phe Leu His Lys Arg Val Leu 210 215 220

Val Asp Lys Asp Ile Asn Leu Leu Ser Ser Ile Glu Arg Ala Leu Glu 225 230 235 240

Leu Pro Thr His Trp Arg Val Gln Met Pro Asn Ala Arg Ser Phe Ile

250

Asp Ala Tyr Lys Arg Arg Pro Asp Met Asn Pro Thr Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Met Val Gln Ala Gln Phe Gln Gln Glu Leu Lys Glu Ala Ser Arg Trp Trp Asn Ser Thr Gly Leu Val His Glu Leu Pro Phe Val Arg Asp Arg Ile Val Glu Cys Tyr Tyr Trp Thr Thr Gly Val Val Glu Arg Arg Glu His Gly Tyr Glu Arg Ile Met Leu Thr Lys Ile Asn Ala Leu Val Thr Thr Ile Asp Asp Val Phe Asp Ile Tyr Gly Thr Leu Glu Glu Leu Gln Leu Phe Thr Thr Ala Ile Gln Arg Trp Asp Ile Glu Ser Met Lys Gln Leu Pro Pro Tyr Met Gln Ile Cys Tyr Leu 375 Ala Leu Phe Asn Phe Val Asn Glu Met Ala Tyr Asp Thr Leu Arg Asp 385 Lys Gly Phe Asn Ser Thr Pro Tyr Leu Arg Lys Ala Trp Val Asp Leu Val Glu Ser Tyr Leu Ile Glu Ala Lys Trp Tyr Tyr Met Gly His Lys 420 425 430 Pro Ser Leu Glu Glu Tyr Met Lys Asn Ser Trp Ile Ser Ile Gly Gly Ile Pro Ile Leu Ser His Leu Phe Phe Arg Leu Thr Asp Ser Ile Glu Glu Glu Asp Ala Glu Ser Met His Lys Tyr His Asp Ile Val Arg Ala Ser Cys Thr Ile Leu Arg Leu Ala Asp Asp Met Gly Thr Ser Leu Asp Glu Val Glu Arg Gly Asp Val Pro Lys Ser Val Gln Cys Tyr Met Asn 505 Glu Lys Asn Ala Ser Glu Glu Glu Ala Arg Glu His Val Arg Ser Leu Ile Asp Gln Thr Trp Lys Met Met Asn Lys Glu Met Met Thr Ser Ser Phe Ser Lys Tyr Phe Val Gln Val Ser Ala Asn Leu Ala Arg Met Ala

Gln Trp Ile Tyr Gln His Glu Ser Asp Gly Phe Gly Met Gln His Ser

565 570 575

Leu Val Asn Lys Met Leu Arg Gly Leu Leu Phe Asp Arg Tyr Glu 580 590

INTERNATIONAL SEARCH REPORT

International application No. PCT/US98/20120

	SSIFICATION OF SUBJECT MATTER Please See Extra Sheet.							
US CI	435/232, 233, 252.3, 320.1, 69.1, 419; 530/350; 536/ o International Patent Classification (IPC) or to both t	23.2, 23.6						
	DS SEARCHED ocumentation searched (classification system followed	by classification symbols)						
	435/232, 233, 252.3, 320.1, 69.1, 419; 530/350; 536/2							
Documentat	ion searched other than minimum documentation to the	extent that such documents are included	in the fields searched					
Electronic d	ata base consulted during the international search (na	me of data base and, where practicable,	search terms used)					
	e Extra Sheet.							
c. Doc	UMENTS CONSIDERED TO BE RELEVANT							
Category*	Citation of document, with indication, where ap	propriate, of the relevant passages	Relevant to claim No.					
CROTEAU, R. et al. Biosynthesis of Monoterpenes: Partial Purification, Characterization, and Mechanism of Action of 1,8-Cineole Synthase. Arch. Biochem. Biophys. 15 February 1994, Vol. 309, No. 1, pages 184-192, see the entire article, especially pages 184 & 187-192.								
Y	LEE, C. C. et al. Generation of cDN. Acid Sequence: Cloning of Urate Oxida Vol. 239, pages 1288-1291, see the er	ase. Science. 11 March 1988,	1, 6-8					
Furth	ner documents are listed in the continuation of Box C	See patent family annex.						
•	ecial categories of cited documents:	"T" later document published after the inte date and not in conflict with the app	ernational filing date or priority					
"A" doe to	cument defining the general state of the art which is not considered be of particular relevance	the principle or theory underlying the	invention					
"E" ORF	lier document published on or after the international filing date	"X" document of particular relevance; the considered novel or cannot be considered.	e claimed invention cannot be red to involve an inventive step					
"L" doc	cument which may throw doubts on priority claim(s) or which is ad to establish the publication date of another citation or other	when the document is taken alone						
spe	icial reason (as specified)	"Y" document of particular relevance; the considered to involve an inventive	step when the document is					
O doc	cument referring to an oral disclosure, use, exhibition or other	combined with one or more other such being obvious to a person skilled in t	documents, such combination					
	nument published prior to the international filing date but later than priority data claimed	*&* document member of the same patent	femily					
	actual completion of the international search	Date of mailing of the international sea	rch report					
05 NOVE	MBER 1998	16DEC 199	38					
Commission Box PCT	nailing address of the ISA/US ner of Patents and Trademarks 1, D.C. 20231	16DEC 199 Authorized Officer TERCHAND SAIDHA	ta					
Facsimile N		Telephone No. (703) 308-0196						

INTERNATIONAL SEARCH REPORT

International application No. PCT/US98/20120

A. CLASSIFICATION OF SUBJECT MATTER: IPC (6):
C12N 1/20, 9/88, 9/90, 15/00; C12P 21/06; C07K 1/00; C07H 21/04
B. FIELDS SEARCHED Electronic data bases consulted (Name of data base and where practicable terms used):
APS. STN Files: caplus, medline, biosis, agricola, wpids, biotechds & biosis. Search terms: bornyl diphosphate synthase, sabinene synthase, 1.8-cineole synthase, etc. Gene and protein sequence data bases searches for Specific SEQ ID NOs.